



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 105144**

**TO: Minh-Tam Davis**  
**Location: CM-1/8A01/8E12**  
**Art Unit: 1642**  
**Tuesday, October 07, 2003**

**Case Serial Number: 09807512**

**From: Toby Port**  
**Location: Biotech-Chem Library**  
**CM1-6A04**  
**Phone: 308-3534**

**[toby.port@uspto.gov](mailto:toby.port@uspto.gov)**

### **Search Notes**

Dear Examiner Davis,

Here are the results of your search.  
Please feel free to contact me if you have any questions.

Toby Port

**THIS PAGE BLANK (USPTO,**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:04:57 ; Search time 18.2405 Seconds  
(without alignments)  
95.721 Million cell updates/sec

Title: US-09-807-512-11  
Perfect score: 50  
Sequence: 1 MLMAQEAFL 11

Scoring table: BLOSUM62  
Gap 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*

```

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	50	100.0	11	21	CTL epitope-1 of h
2	50	100.0	11	23	Tumour antigen TAG
3	50	100.0	30	23	Human NYSOLb seqm
4	50	100.0	58	20	Human cancer antig
5	50	100.0	58	21	Human tumour antig
6	50	100.0	58	23	Human NYSOLb cons
7	50	100.0	109	21	Human CTL-recognis
8	50	100.0	3541	23	Human melanoma spe
9	46	92.0	10	23	Human CAMEL class

10	45	90.0	10	21	AA170856	CTL epitope-2 of h
11	34	68.0	222	23	ABF73965	Candida albicans e
12	34	68.0	277	23	ABU65209	Human NOV121a prot
13	33	66.0	272	22	AAU34739	E. coli cellular p
14	33	66.0	320	21	AG222016	Arabidopsis thalia
15	33	66.0	761	22	AA471134	CDIFF-15, incyte I
16	33	66.0	891	23	AAE22231	Human schlafen C p
17	32	64.0	63	23	ABF34124	Human ORF3097 prot
18	32	64.0	92	20	AA113124	Human secreted pro
19	32	64.0	105	23	AB447977	Listeria monocytog
20	32	64.0	215	23	AB934441	Herbicidally activ
21	32	64.0	257	24	AB47519	Breast cancer asso
22	32	64.0	363	23	AB55383	Lactococcus lactis
23	32	64.0	463	21	AAV50930	Human fetal brain
24	32	64.0	463	21	AAV44276	Human nucleic acid
25	32	64.0	552	22	AB85484	pfam consensus seq
26	32	64.0	873	23	ABF30006	Streptococcus poly
27	31	62.0	83	22	AAU55483	Propionibacterium
28	31	62.0	100	22	AAU16254	Human novel secret
29	31	62.0	100	24	ABU55323	Human novel poly
30	31	62.0	103	22	AAU19458	Human diagnostic a
31	31	62.0	124	21	AB27713	Human secreted pro
32	31	62.0	143	23	AB93600	Herbicidally activ
33	31	62.0	290	23	AB89676	Human polypeptide
34	31	62.0	395	23	ABJ11257	Yeast selected int
35	31	62.0	395	23	ABG77355	Selected interacti
36	31	62.0	398	22	AB448047	Signal transductio
37	31	62.0	398	22	AB48048	Signal transductio
38	31	62.0	576	22	AA40397	Human polypeptide
39	31	62.0	619	24	AB82485	M. sexta acetylcho
40	31	62.0	636	20	AAV34710	Amino acid sequenc
41	31	62.0	780	22	AAE03802	Dictyostelium disc
42	31	62.0	914	21	AG49883	Arabidopsis thalia
43	31	62.0	924	21	AG49882	Arabidopsis thalia
44	31	62.0	1176	21	AG49881	Arabidopsis thalia
45	31	62.0	3435	22	AB59499	Drosophila melanog

#### ALIGNMENTS

##### RESULT 1

AA170855

ID AA170855 standard; peptide; 11 AA.

XX

AC AA170855;

XX

DT 31-JUL-2000 (first entry)

XX

DE CTL epitope-1 of human CAMEL protein.

XX

KW CAMEL; CTL-recognised Antigen on MELanoma; cytotoxic T lymphocyte; CTL;  
KW tumour-associated antigen; LAGE-1; NY-ESO-1; anticancer; melanoma;  
KW human; cancer; immunotherapy; immunogenic peptide; immune response.

XX

OS Homo sapiens.

XX

PN WO200023584-A1.

XX

PD 27-APR-2000.

XX

PF 15-OCT-1999; 99WO-EP07832.

XX

PR 16-OCT-1998; 98EP-0119583.

XX

PA (BOEH ) BOEHRINGER INGELHEIM INT GMBH.  
PA (UYHO-) UNIV HOSPITAL LEIDEN.

XX

PI Schrier PI, Aarnoudse CA, Heider K, Klade C;

XX

DR WPI; 2000-339685/29.

XX

PT Tumor-associated antigen useful for cancer immunotherapy is encoded by

PT the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA -  
XX Claim 4; Page 34; 73pp; English.  
PS  
CC The present sequence is an immunogenic peptide of human tumour-associated  
CC antigen CAMEL (Cytotoxic T lymphocytes (CTL)-recognised Antigen on  
CC MELANOMA). This peptide is a CTL epitope, that has the ability to elicit  
CC a CTL response. It corresponds to residues 1-11 of the CAMEL protein.  
CC CAMEL protein is encoded by the LAGE-1 gene, a tumour-specific antigen.  
CC It is different from the LAGE-1 protein, since it is translated from a  
CC different open reading frame (ORF-1). It shows strong homology with  
CC NY-ESO-1, a melanoma specific tumour antigen. The tumour-associated  
CC antigen displayed on melanoma cells is recognised by cytotoxic T  
CC lymphocytes. CAMEL is expressed in tumour cell lines, tumour tissues  
CC (e.g. breast and lung) and in restricted number of healthy tissues. This  
CC sequence has anticancer activity. CAMEL tumour antigen and immunogenic  
CC peptides derived from it are useful for cancer immunotherapy. They have  
CC the potential to induce an immune response, by eliciting a CTL response.  
CC The DNA molecule is used to construct recombinant or fusion proteins.  
XX  
SQ Sequence 11 AA;  
Query Match 100.0%; Score 50; DB 21; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0039;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MLMAQEALAFLL 11  
DB 1 MLMAQEALAFLL 11  
RESULT 2  
ABG66804  
ID ABG66804 standard; Peptide; 11 AA.  
XX  
AC ABG66804;  
XX  
DT 24-SEP-2002 (first entry)  
XX  
DE Tumour antigen LAGE/CAMEL, HLA-A2 epitope.  
XX  
KW Beta-2 microglobulin; beta-2m; cytotoxic T lymphocyte; CTL; HLA;  
KW human leukocyte antigen; fusion protein; epitope; cytostatic; tumour;  
KW gastrointestinal tumour; colorectal cancer; gastro-oesophageal cancer;  
KW liver cancer; biliary tract cancer; pancreatic cancer; vaccine;  
KW prostatic cancer; testicular cancer; lung cancer; breast cancer;  
KW malignant melanoma; mesothelioma; brain tumour; ovarian cancer;  
KW uterine cancer; cervical cancer; head and neck cancer; bladder cancer;  
KW Kaposi's sarcoma; renal carcinoma; leukaemia; lymphoma;  
KW acquired immunodeficiency syndrome; AIDS-related lymphoma.  
XX  
OS Homo sapiens.  
XX  
XX WO200236146-A2.  
PN  
XX 10-MAY-2002.  
XX  
XX 01-NOV-2001; 2001WO-GB04844.  
XX  
XX 02-NOV-2000; 2000GB-0026912.  
XX  
XX (ISIS-) ISIS INNOVATION LTD.  
XX  
XX Tafuro S, Meier U, McMichael AJ, Bell JI, Layton G, Hunter M;  
PI  
XX WPI; 2002-508108/54.  
DR  
XX New polynucleotide capable of expressing an epitope-beta2m fusion  
XX protein useful for generating cytotoxic T lymphocyte responses against  
PT a tumour and in restoring antigen presentation in the tumour of a host  
PT  
XX  
XX Disclosure; Page 25; 46pp; English.  
PS

XX The invention relates to a new polynucleotide capable of expressing an  
CC epitope-beta2m fusion protein useful for generating cytotoxic T  
CC lymphocyte (CTL) responses against a tumour or in restoring antigen  
CC presentation in the tumour of a host. Also included are a polynucleotide  
CC capable of expressing an epitope-beta2m fusion protein in combination  
CC with a vaccination agent that stimulates a CTL response against the  
CC epitope of the fusion protein for simultaneous, separate or sequential  
CC use in the treatment of cancer and a method of treating a tumour by  
CC administering a capable of expressing an epitope-beta2m fusion protein,  
CC and optionally a vaccination agent that stimulates a CTL response against  
CC the epitope of the fusion protein. The polynucleotide is useful for  
CC generating CTL responses against tumours, for restoring antigen  
CC presentation in the tumour, and subsequently for treating cancers, such  
CC as gastrointestinal tumour, prostatic, testicular, lung or breast cancer,  
CC malignant melanoma, mesothelioma, brain tumour, ovarian cancer, uterine  
CC cancer including cervical cancer, cancer of the head and neck, bladder  
CC cancer, Kaposi's sarcoma, AIDS (acquired immunodeficiency syndrome).  
CC related Kaposi's sarcoma, sarcomas, osteosarcoma, renal carcinoma, and  
CC haematopoietic malignant tumours such as leukaemia and lymphoma.  
CC The epitope is an HLA (human leukocyte antigen) peptide derived from a  
CC viral or tumour antigen. The present sequence is a tumour HLA epitope  
CC used in the fusion proteins of the invention.  
XX  
SQ Sequence 11 AA;  
Query Match 100.0%; Score 50; DB 23; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0039;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MLMAQEALAFLL 11  
DB 1 MLMAQEALAFLL 11  
RESULT 3  
AAU85114  
ID AAU85114 standard; Peptide; 30 AA.  
XX  
AC AAU85114;  
XX  
DT 08-MAY-2002 (first entry)  
XX  
DE Human NYNSO1b segment 1.  
XX  
KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;  
KW viral infection; human immunodeficiency virus; melanoma;  
KW bacterial infection; Salmonella; Legionella; parasitic infection;  
KW Trypanosoma; Toxoplasma; Giardia.  
XX  
OS Homo sapiens.  
XX  
XX WO200190197-A1.  
PN  
XX 29-NOV-2001.  
XX  
XX 25-MAY-2001; 2001WO-AU00622.  
XX  
XX 26-MAY-2000; 2000AU-0007761.  
XX  
XX (AUSU) UNIV AUSTRALIAN NAT.  
XX  
XX Thomson SA, Ramshaw IA;  
PI  
XX WPI; 2002-147575/19.  
DR  
XX N-PSDB; ABK36934.  
DR  
XX New synthetic polypeptides having several different segments of at  
PT least one parent polypeptide linked together differently compared to  
PT the linkage in the parent polypeptide, for inducing immune response  
PT against a pathogen or cancer  
XX  
XX Example 3; Fig 27; 364pp; English.  
PS



CC tumour antigen, identified by screening an esophagus carcinoma cDNA  
CC library. This protein is derived from open reading frame (ORF)-2,  
CC that differs from the CAMEL protein (Cytotoxic T lymphocytes (CTL)-  
CC recognised antigen on MELANOMA), a tumour-associated antigen, by only its  
CC last 5 amino acids. It contains epitopes of tumour specific T-cells.  
CC NY-ESO-1 is expressed in different tumour types, but not in healthy  
CC tissues except in testis. The tumour-associated antigen displayed on  
CC melanoma cells is recognised by cytotoxic T lymphocytes. This sequence  
CC has anticancer activity. CAMEL tumour antigen and immunogenic peptides  
CC derived from it are useful for cancer immunotherapy. They have the  
CC potential to induce an immune response, by eliciting a CTL response.  
CC The DNA molecule is used for the construction of recombinant or fusion  
CC proteins.  
CC XX  
SQ Sequence 58 AA;  
Query Match 100.0%; Score 50; DB 21; Length 58;  
Best Local Similarity 100.0%; Pred. No. 0.022;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MLMQAQALAF 11  
DB 1 MLMQAQALAF 11  
RESULT 6  
AAU84819  
ID AAU84819 standard; Protein: 58 AA.  
XX  
AC AAU84819;  
XX  
DT 08-MAY-2002 (first entry)  
XX  
DE Human NYSOLB consensus sequence.  
XX  
XX Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;  
KW viral infection; human immunodeficiency virus; melanoma;  
KW bacterial infection; Salmonella; Legionella; parasitic infection;  
KW Trypanosoma; Toxoplasma; Giardia.  
XX  
OS Homo sapiens.  
XX  
XX WO200190197-A1.  
XX  
XX 29-NOV-2001.  
XX  
XX 25-MAY-2001; 2001WO-AU00622.  
XX  
XX 26-MAY-2000; 2000AU-0007761.  
XX  
XX (AUSU ) UNIV AUSTRALIAN NAT.  
XX  
XX Thomson SA, Ramshaw IA;  
XX  
XX WPI; 2002-147575/19.  
XX  
XX New synthetic polypeptides having several different segments of at  
XX least one parent polypeptide linked together differently compared to  
XX the linkage in the parent polypeptide, for inducing immune response  
XX against a pathogen or cancer  
XX  
XX Example 3; Fig 27; 364pp; English.  
XX  
XX The invention relates to a new synthetic polypeptide (1) comprising  
XX several different segments of at least one parent polypeptide linked  
XX together in a different relationship relative to their linkage in the  
XX parent polypeptide to impede, abrogate or otherwise alter at least one  
XX function associated with the parent polypeptide and for inducing an  
XX immune response against a pathogen or cancer. Also included are a  
XX synthetic polynucleotide encoding and a computer system for  
XX designing the synthetic polypeptides. The synthetic polypeptides and  
XX polynucleotides are referred to as a Savine. The synthetic polypeptide is  
XX useful for modulating immune responses preferably directed against a

CC pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix,  
CC colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone  
CC liver, oesophagus, brain, testicle, uterus), as potentiating agents.  
CC Compositions comprising the polypeptide may be used in the treatment or  
CC prophylaxis against viral (such as infections caused by HIV (human  
CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis  
CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial  
CC (e.g., infections caused by Neisseria, Meningococcus, Haemophilus,  
CC Salmonella, Streptococcus, Legionella and Mycobacterium or parasitic  
CC infections caused by Plasmodium, Schistosoma, Leishmania,  
CC Trypanosoma, Toxoplasma and Giardia) infections. The present  
CC sequence is a consensus sequence for a parent protein used to design a  
CC savine of the invention.  
CC XX  
SQ Sequence 58 AA;  
Query Match 100.0%; Score 50; DB 23; Length 58;  
Best Local Similarity 100.0%; Pred. No. 0.022;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MLMQAQALAF 11  
DB 1 MLMQAQALAF 11  
RESULT 7  
AAV70854  
ID AAV70854 standard; Protein: 109 AA.  
XX  
AC AAV70854;  
XX  
DT 31-JUL-2000 (first entry)  
XX  
DE Human CTL-recognised Antigen on MELANOMA (CAMEL) protein.  
XX  
XX CAMEL; CTL-recognised Antigen on MELANOMA; cytotoxic T lymphocyte; CTL;  
KW tumour-associated antigen; LAGE-1; NY-ESO-1; anticancer; melanoma;  
KW cancer; immunotherapy; immunogenic peptide; immune response.  
XX  
OS Homo sapiens.  
XX  
XX Key  
XX Location/Qualifiers  
XX 1..11  
XX /label= CTL epitope  
XX /note= "Immunogenic peptide with ability to elicit a  
XX CTL response"  
XX 2..10  
XX /label= Immunogenic peptide  
XX /note= "Specific for HLA-A3"  
XX 2..11  
XX /label= CTL epitope  
XX /note= "Immunogenic peptide with ability to elicit a  
XX CTL response"  
XX 10..18  
XX /label= CAMEL\_10  
XX /note= "Specific for HLA-A2"  
XX 16..25  
XX /label= CAMEL\_16  
XX /note= "Specific for HLA-A2"  
XX 17..25  
XX /label= CAMEL\_17  
XX /note= "Specific for HLA-A2"  
XX 51..59  
XX /label= Immunogenic peptide  
XX /note= "Specific for HLA-A3101"  
XX 101..109  
XX /label= Immunogenic peptide  
XX  
XX WO200023584-A1.  
XX  
XX 27-APR-2000.  
XX  
XX 15-OCT-1999; 99WO-EP07832.



XX WPI; 2002-627577/67.  
 XX Novel composition for treating a disease in an animal, comprises an  
 PT immune effector cell and cell penetrating peptide associated with an  
 PT antigen or antibody -  
 XX  
 PS Disclosure; Page 17; 61pp; English.  
 XX  
 CC The invention relates to a composition (I) comprising an immune effector  
 CC cell and a cell penetrating peptide (CPP) associated with an antigen or  
 CC antibody. Also included are (1) a vaccine comprising (I), CPP  
 CC associated with an antigen, and a pharmaceutically acceptable carrier  
 CC and (2) preparing a composition for a disease, by providing (I)  
 CC and CPP associated with an antigen for disease, and introducing the  
 CC antigen-associated CPP to (I), where antigen enters into the cell.  
 CC The antigens are, for example, tumour antigen derived epitopes  
 CC recognised by tumour infiltrating lymphocytes (TIL) of HLA (human  
 CC leukocyte antigen) class I or II. The composition is useful for enhancing  
 CC immunity in an animal to a disease, by administering a mature dendritic  
 CC cell comprising CPP associated with an antigen to disease, to the animal,  
 CC such that following the administration, animal is protected from disease,  
 CC where the animal comprises both CD4+ and CD8+ T cells. It is also useful  
 CC for treating a disease (e.g. cancer, tumour, melanoma, thymoma,  
 CC lymphoma, sarcoma, lung cancer, non-Hodgkin's lymphoma, leukaemia,  
 CC Hodgkin's lymphoma, uterine cancer, cervical cancer, bladder cancer,  
 CC kidney cancer, adenocarcinoma, breast cancer, prostate cancer,  
 CC ovarian cancer and pancreatic cancer). The animal is further subjected to  
 CC a cancer treatment including surgery, radiation, chemotherapy or gene  
 CC therapy. The administration of (I), preferably dendritic cell is prior  
 CC to, subsequent to or concurrent with, the cancer treatment. The present  
 CC sequence is a tumour antigen derived epitope for inclusion in the  
 CC composition of the invention.  
 XX  
 SQ Sequence 10 AA;  
 Query Match 92.0%; Score 46; DB 23; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.021;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MLMAQEALAF 10  
 | | | | |  
 DB 1 MLMAQEALAF 10  
 RESULT 10  
 AAY70856  
 ID AAY70856 standard; peptide; 10 AA.  
 XX  
 AC AAY70856;  
 XX  
 DT 31-JUL-2000 (first entry)  
 XX  
 DE CTL epitope-2 of human CAMEL protein.  
 XX  
 KW CAMEL; CTL-recognised Antigen on MELanoma; cytotoxic T lymphocyte; CTL;  
 KW tumour-associated antigen; LAGE-1; NY-ESO-1; anticancer; melanoma;  
 KW human; cancer; immunotherapy; immunogenic peptide; immune response.  
 XX Homo sapiens.  
 OS  
 PN WO200023584-A1.  
 XX  
 PD 27-APR-2000.  
 XX  
 PF 15-OCT-1999; 99WO-EP07832.  
 XX  
 PR 16-OCT-1998; 98EP-0119583.  
 XX  
 PA (BOEH ) BOEHRINGER INGELHEIM INT GMBH.  
 PA (UYHO-) UNIV HOSPITAL LEIDEN.  
 XX  
 PI Schrier PI, Aarnoudse CA, Heider K, Klade C;

XX WPI; 2000-339685/29.  
 XX Tumor-associated antigen useful for cancer immunotherapy is encoded by  
 PT the open reading frame of LAGE-1 (a tumor-specific antigen) CDNA -  
 XX  
 PS Claim 5; Page 34; 73pp; English.  
 XX  
 CC The present sequence is an immunogenic peptide of human tumour-associated  
 CC antigen CAMEL (Cytotoxic T lymphocytes (CTL)-recognised Antigen on  
 CC MELanoma). This peptide is a CTL epitope, that has the ability to elicit  
 CC a CTL response. It corresponds to residues 2-11 of the CAMEL protein.  
 CC CAMEL protein is encoded by the LAGE-1 gene, a tumour-specific antigen.  
 CC It is different from the LAGE-1 protein, since it is translated from a  
 CC different open reading frame (ORF-1). It shows strong homology with  
 CC NY-ESO-1, a melanoma specific tumour antigen. The tumour-associated  
 CC antigen displayed on melanoma cells is recognised by cytotoxic T  
 CC lymphocytes. CAMEL is expressed in tumour cell lines, tumour tissues  
 CC (e.g. breast and lung) and in restricted number of healthy tissues. This  
 CC sequence has anticancer activity. CAMEL tumour antigen and immunogenic  
 CC peptides derived from it are useful for cancer immunotherapy. They have  
 CC the potential to induce an immune response, by eliciting a CTL response.  
 CC The DNA molecule is used to construct recombinant or fusion proteins.  
 XX  
 SQ Sequence 10 AA;  
 Query Match 90.0%; Score 45; DB 21; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.033;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 LMAQEALAF 11  
 | | | | |  
 DB 1 LMAQEALAF 10  
 RESULT 11  
 ABP73965  
 ID ABP73965 standard; Protein; 222 AA.  
 XX  
 AC ABP73965;  
 XX  
 DT 30-JAN-2003 (first entry)  
 XX  
 DE Candida albicans essential protein SEQ ID NO 7802.  
 XX  
 KW Fungus; yeast; tetracyclin; promoter; GRACE strain; biosynthesis;  
 KW signal transduction; DNA replication; cell division; growth;  
 KW proliferation; Candida albicans; fungicide; antifungal.  
 XX  
 OS Candida albicans.  
 PN WO200253728-A2.  
 XX  
 PD 11-JUL-2002.  
 XX  
 PF 26-DEC-2001; 2001WO-US49486.  
 XX  
 PR 29-DEC-2000; 2000US-259128P.  
 PR 20-FEB-2001; 2001US-0792024.  
 PR 22-AUG-2001; 2001US-314050P.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.  
 XX  
 PI Roemer T, Jiang B, Boone C, Bussey H, Ohlsen KL;  
 XX  
 DR WPI; 2002-566694/60.  
 DR N-PSDB; ABZ32515.  
 XX  
 CC Constructing strains for identifying gene products as effective targets  
 PT for therapeutic intervention, by inactivating in the strain one allele  
 PT of a gene and placing other allele of the gene under conditional  
 PT expression -  
 XX

PS Claim 44; SEQ ID NO 7802; 167pp + Sequence Listing; English.

XX The invention relates to constructing (M1) a strain of diploid fungal

CC cells in which both alleles of a gene are modified, comprising modifying

CC one allele by insertion or replacement by a cassette having an

CC expressible selectable marker and modifying other allele by

CC recombination, of a promoter replacement fragment with a heterologous

CC promoter, so that expression of the second allele is regulated by the

CC promoter. (M1) is useful for constructing a strain of diploid fungal

CC cells in which both alleles of a gene are modified. The diploid fungal

CC cells having both alleles modified are useful for identifying a gene that

CC is essential to the survival or growth of a fungus, a gene that

CC contributes to the virulence and/or pathogenicity of a fungus, a gene

CC that contributes to the resistance of a diploid fungus to an antifungal

CC agent, an antifungal agent that inhibits the growth of a diploid fungus

CC and for identifying a therapeutic agent for treatment of a mammalian

CC disease. (M1) is useful for identifying a compound which modulates the

CC activity of a gene product, preferably enzymatic activity, carbon

CC compound catabolism, biosynthetic, transporter, transcriptional,

CC translational, signal transduction, DNA replication and cell division

CC activity. The method is useful for identifying a compound having the

CC ability to inhibit growth or proliferation of C. albicans cells and for

CC treating infection by C. albicans. The present sequence is that of an

CC essential Candida albicans protein used in the method of the invention.

CC Note: The sequence data for this patent is not represented in the printed

CC specification but is based on sequence information supplied to Derwent by

CC the European Patent Office.

XX

XX Sequence 222 AA;

Query Match 68.0%; Score 34; DB 23; Length 222;

Best Local Similarity 80.0%; Pred. No. 1e+02;

Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LMAQELAEFL 11

Db 122 LMAQEADAEV 131

||||| 11:

RESULT 12

ABU65209

ID ABU65209 standard; Protein; 577 AA.

XX

AC ABU65209;

XX

DT 20-MAY-2003 (first entry)

XX

DE Human NOV121a protein.

XX

XX NOVX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;

KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;

KW human.

XX

OS Homo sapiens.

XX

PN WO200272757-A2.

XX

XX 19-SEP-2002.

PD

XX

XX 08-MAR-2002; 2002WO-US06908.

PF

XX

PR 08-MAR-2001; 2001US-274101P.

PR

PR 08-MAR-2001; 2001US-274194P.

PR

PR 08-MAR-2001; 2001US-274281P.

PR

PR 08-MAR-2001; 2001US-274322P.

PR

PR 09-MAR-2001; 2001US-274849P.

PR

PR 12-MAR-2001; 2001US-275235P.

PR

PR 13-MAR-2001; 2001US-275578P.

PR

PR 13-MAR-2001; 2001US-275579P.

PR

PR 13-MAR-2001; 2001US-275601P.

PR

PR 14-MAR-2001; 2001US-276000P.

PR

PR 16-MAR-2001; 2001US-276776P.

PR

PR 19-MAR-2001; 2001US-276994P.

PR 20-MAR-2001; 2001US-277239P.

PR

PR 20-MAR-2001; 2001US-277321P.

PR

PR 20-MAR-2001; 2001US-277327P.

PR

PR 21-MAR-2001; 2001US-277791P.

PR

PR 22-MAR-2001; 2001US-277833P.

PR

PR 23-MAR-2001; 2001US-278152P.

PR

PR 26-MAR-2001; 2001US-278894P.

PR

PR 27-MAR-2001; 2001US-278999P.

PR

PR 27-MAR-2001; 2001US-279036P.

PR

PR 28-MAR-2001; 2001US-279344P.

PR

PR 30-MAR-2001; 2001US-277338P.

PR

PR 30-MAR-2001; 2001US-279995P.

PR

PR 30-MAR-2001; 2001US-280233P.

PR

PR 02-APR-2001; 2001US-280802P.

PR

PR 02-APR-2001; 2001US-280822P.

PR

PR 02-APR-2001; 2001US-280900P.

PR

PR 04-APR-2001; 2001US-281194P.

PR

PR 13-APR-2001; 2001US-283675P.

PR

PR 30-APR-2001; 2001US-287424P.

PR

PR 02-MAY-2001; 2001US-288066P.

PR

PR 03-MAY-2001; 2001US-288342P.

PR

PR 03-MAY-2001; 2001US-288528P.

PR

PR 13-MAY-2001; 2001US-291190P.

PR

PR 16-MAY-2001; 2001US-291099P.

PR

PR 16-MAY-2001; 2001US-291240P.

PR

PR 30-MAY-2001; 2001US-294485P.

PR

PR 31-MAY-2001; 2001US-294889P.

PR

PR 31-MAY-2001; 2001US-294899P.

PR

PR 18-JUN-2001; 2001US-299027P.

PR

PR 19-JUN-2001; 2001US-299303P.

PR

PR 19-JUN-2001; 2001US-299310P.

PR

PR 10-JUL-2001; 2001US-304354P.

PR

PR 31-JUL-2001; 2001US-309198P.

PR

PR 16-AUG-2001; 2001US-312903P.

PR

PR 10-SEP-2001; 2001US-318462P.

PR

PR 12-SEP-2001; 2001US-318770P.

PR

PR 27-SEP-2001; 2001US-325430P.

PR

PR 27-SEP-2001; 2001US-325681P.

PR

PR 18-OCT-2001; 2001US-325681P.

PR

PR 31-OCT-2001; 2001US-335301P.

PR

PR 14-NOV-2001; 2001US-332172P.

PR

PR 14-NOV-2001; 2001US-332271P.

PR

PR 14-NOV-2001; 2001US-332272P.

PR

PR 14-NOV-2001; 2001US-333184P.

PR

PR 14-NOV-2001; 2001US-333272P.

PR

PR 21-NOV-2001; 2001US-332094P.

PR

PR 03-DEC-2001; 2001US-337426P.

PR

PR 03-DEC-2001; 2001US-338092P.

PR

PR 04-DEC-2001; 2001US-337185P.

PR

PR 03-JAN-2002; 2002US-345705P.

PR

PR 07-MAR-2002; 2002US-0092900.

XX

XX

PA (CURA-) CURAGEN CORP.

XX

XX

PI Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;

PI Zerhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;

PI Patturajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;

PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y;

PI Anderson D, Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H;

PI Alsobrook JP, Lepley DM, Rieger DK;

XX

XX WPI: 2002-723332/78.

DR

DR N-PSDB; ABX971176.

XX

XX

PT NOVX polypeptides and polynucleotides, useful for preventing or

PT treating a disorder associated with aberrant NOVX expression or

PT activity e.g., cancer, hypertension, atherosclerosis, cardiomyopathy or

PT bronchial asthma

XX

PS Claim 1; Page 455; 1103pp; English.

XX

CC This invention describes novel human NOVX polypeptides which have

CC cytostatic, cardiant, antiarteriosclerotic, antiasthmatic and

Tue Oct 7 15:52:17 2003

CC for homologous nucleic acids which are required for cell proliferation in  
CC a wide variety of organisms. The present sequence represents an  
CC essential prokaryotic cellular proliferation protein. part  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 272 AA;

Query Match 66.0%; Score 33; DB 22; Length 272;  
Best Local Similarity 72.7%; Pred. No. 2e+02; 0; Gaps 0;  
Matches 8; Conservative 1; Mismatches 2; Indels 0;  
QY 1 MLMAQEALAF 11  
Db 240 MLVAQAHAFL 250  
||:|||||  
||:|||||

RESULT 14  
AAG22016  
ID AAG22016 standard; Protein; 320 AA.  
XX  
AC AAG22016;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 24785.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.

CC hypotensive activity. Pharmaceutical compositions comprising the NOVX  
CC proteins or nucleic acid molecules or NOVX antibodies are useful for  
CC preventing or treating a disorder associated with aberrant NOVX  
CC expression or activity e.g. cancer, hypertension, atherosclerosis,  
CC cardiomyopathy or bronchial asthma. The products of the invention can  
CC be used for gene therapy or in a vaccine. ABU65041-ABU65218 represent  
CC the NOVX polypeptides encoded by ABX97008-ABX97185.  
XX  
SQ Sequence 577 AA;

Query Match 68.0%; Score 34; DB 23; Length 577;  
Best Local Similarity 70.0%; Pred. No. 2.7e+02;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 2 LMAQEALAF 11  
Db 260 LMLKALAFI 269  
||:|||||  
||:|||||

RESULT 13  
AAU34739  
ID AAU34739 standard; Protein; 272 AA.  
XX  
AC AAU34739;  
XX  
DT 14-FEB-2002 (first entry)  
XX  
DE E. coli cellular proliferation protein #320.  
XX  
KW Antisense; prokaryotic cellular proliferation protein;  
KW antibiotic; antibacterial; drug design.  
XX  
OS Escherichia coli.  
XX  
PN WO200170955-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 21-MAR-2001; 2001WO-US09180.  
XX  
PR 21-MAR-2000; 2000US-191078P.  
PR 23-MAY-2000; 2000US-206848P.  
PR 26-MAY-2000; 2000US-207727P.  
PR 23-OCT-2000; 2000US-242578P.  
PR 27-NOV-2000; 2000US-253625P.  
PR 22-DEC-2000; 2000US-257931P.  
PR 16-FEB-2001; 2001US-269308P.

XX  
PA (ELIT-) ELITRA PHARM INC.  
XX  
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
PI Yamamoto RT, Xu HH;  
XX  
DR WPI: 2001-611495/70.  
DR N-PSDB; RAS52598.  
XX  
PT New polynucleotides for the identification and development of  
PT antibiotics, comprise sequences of antisense nucleic acids -  
XX  
PS Example 3; Seq ID No 10332; 51lpp; English.

XX The invention relates to antisense inhibitors of genes essential to  
XX prokaryotic cellular proliferation, their use in identifying the  
XX genes, their use in the discovery of novel antibiotics, the essential  
XX genes themselves and the encoded proteins. The prokaryotes used are  
XX Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella  
XX pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The  
XX invention is also useful for the identification of potential new targets  
XX for antibiotic development. The antisense nucleic acids can also be used  
XX to identify proteins used in proliferation, to express these proteins,  
XX and to obtain antibodies capable of binding to the expressed proteins.  
XX The proteins can be used to screen compounds in rational drug discovery  
XX programmes. The antisense nucleic acid sequence is also useful to screen





PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 66.0%; Score 33; DB 21; Length 320;  
Best Local Similarity 77.8%; Pred. No. 2.3e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 MAQEALAF 11  
DB 1 MQEALSFL 9

RESULT 15  
AAB47134  
ID AAB47134 standard; Protein; 761 AA.

AC AAB47134;

DT 04-JUN-2001 (first entry)

DE CDIFF-15, Incyte ID No. 3478571CD1.

Human; cell differentiation; CDIFF; agonist; antagonist; epilepsy;  
cell proliferation; Alzheimer's disease; schizophrenic disorder;  
arteriosclerosis; cancer; atherosclerosis; diabetes mellitus.

OS Homo sapiens.

Key Location/Qualifiers

FT Modified-site	67	/note= "Potential phosphorylation site"
FT Modified-site	84	/note= "Potential phosphorylation site"
FT Modified-site	123	/note= "Potential phosphorylation site"
FT Modified-site	148	/note= "Potential phosphorylation site"
FT Modified-site	169	/note= "Potential glycosylation site"
FT Modified-site	193	/note= "Potential phosphorylation site"
FT Modified-site	250	/note= "Potential phosphorylation site"
FT Modified-site	269	/note= "Potential phosphorylation site"
FT Modified-site	355	/note= "Potential phosphorylation site"
FT Modified-site	382	/note= "Potential phosphorylation site"
FT Modified-site	457	/note= "Potential phosphorylation site"
FT Modified-site	473	/note= "Potential phosphorylation site"
FT Modified-site	485	/note= "Potential phosphorylation site"
FT Modified-site	499	/note= "Potential phosphorylation site"
FT Modified-site	557	/note= "Potential phosphorylation site"
FT Modified-site	582	/note= "Potential phosphorylation site"
FT Modified-site	622	/note= "Potential phosphorylation site"
FT Modified-site	631	/note= "Potential phosphorylation site"
FT Modified-site	662	/note= "Potential phosphorylation site"
FT Modified-site	671	/note= "Potential phosphorylation site"

PN WO200119860-A2.

XX PD

XX 22-MAR-2001.

PF 14-SEP-2000; 2000WO-US25435.

PR 15-SEP-1999; 99US-0154140.

PR 06-DEC-1999; 99US-0169155.

XX (INCY-) INCYTE GENOMICS INC.

XX Tang YT, Hillman JL, Yue H, Reddy R, Lal P, Shah P, Azimzai Y;

PI Baughn MR, Lu DAM, Bandman O, Shih LL, Patterson C;

XX WPI; 2001-211447/21.

DR N-PSDB; AAC85556.

XX Isolated polypeptides and polynucleotides involved in cell  
PT differentiation are used for treatment, prevention and diagnosis of  
PT cell proliferative, developmental and neurological disorders e.g.  
PT cancer and Alzheimer's disease -

XX Claim 1; Page 107-109; 137pp; English.

XX The sequences given in AAB47123-50 show human polypeptides involved  
CC in cell differentiation (CDIFF). CDIFF polypeptides and agonists of  
CC these are used to treat a disease or condition associated with  
CC decreased expression of functional CDIFF. An antagonist of CDIFF is  
CC used to treat a disease or condition associated with over expression  
CC of functional CDIFF. CDIFF polypeptides may be used for the treatment,  
CC prevention and diagnosis of cell proliferative, developmental and  
CC neurological disorders, such as Alzheimer's disease, schizophrenic  
CC disorders, arteriosclerosis, cancer, atherosclerosis, diabetes mellitus  
CC and epilepsy. The CDIFF-15 sequence is homologous to Mus musculus  
CC schlafen3.

XX Sequence 761 AA;

Query Match 66.0%; Score 33; DB 22; Length 761;

Best Local Similarity 87.5%; Pred. No. 5.7e+02;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 AQEALAF 11

DB 130 SQEALAF 137

Search completed: October 7, 2003, 13:25:29

Job time : 19.2405 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:18:33 ; Search time 18.3797 Seconds  
(without alignments)  
154.441 Million cell updates/sec

Title: US-09-807-512-11  
Perfect score: 50  
Sequence: 1 MLMQAEALAF 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL\_23.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	100.0	58	095146	095146 homo sapien
2	50	100.0	109	095987	095987 homo sapien
3	35	70.0	308	Q91624	Q91624 pseudomonas
4	35	70.0	654	Q8DGN6	Q8DGN6 synchococcc
5	34	68.0	310	Q9X330	Q9X330 bacillus an
6	34	68.0	388	Q92DV2	Q92DV2 listeria in
7	34	68.0	462	Q94GH7	Q94GH7 oryza sativ
8	34	68.0	837	Q9LZC7	Q9LZC7 arabidopsis
9	34	68.0	896	Q9FFN1	Q9FFN1 arabidopsis
10	33	66.0	272	Q8FD18	Q8FD18 escherichia
11	33	66.0	318	Q8S8E4	Q8S8E4 arabidopsis
12	33	66.0	320	Q8LF84	Q8LF84 arabidopsis
13	33	66.0	333	Q9SJ50	Q9SJ50 arabidopsis
14	33	66.0	338	Q8WU54	Q8WU54 homo sapien
15	33	66.0	665	Q9A7F1	Q9A7F1 caulobacter
16	33	66.0	782	Q96A82	Q96A82 homo sapien

17	32	64.0	105	9	Q9T1A2	Q9T1A2 bacterioph
18	32	64.0	105	16	Q8V4Z7	Q8V4Z7 listeria mo
19	32	64.0	196	4	Q9NOF0	Q9NOF0 homo sapien
20	32	64.0	204	2	Q93MU9	Q93MU9 streptococ
21	32	64.0	205	16	Q9KVK7	Q9KVK7 vibrio chol
22	32	64.0	215	10	Q9LYH8	Q9LYH8 arabidopsis
23	32	64.0	257	4	Q9HC13	Q9HC13 homo sapien
24	32	64.0	373	16	Q8FP51	Q8FP51 corynebacte
25	32	64.0	458	4	Q8WVP3	Q8WVP3 homo sapien
26	32	64.0	460	11	Q9CWH5	Q9CWH5 mus musculu
27	32	64.0	592	2	Q9S433	Q9S433 myxococcus
28	32	64.0	602	16	Q99QL3	Q99QL3 streptomyce
29	32	64.0	623	17	Q9PTS8	Q9PTS8 methanosarc
30	32	64.0	873	16	Q9A1H3	Q9A1H3 streptococcc
31	32	64.0	873	16	Q8P2N1	Q8P2N1 streptococcc
32	32	64.0	873	16	Q8K8N0	Q8K8N0 streptococcc
33	32	64.0	1316	3	Q8TFG6	Q8TFG6 schizosacch
34	32	64.0	2692	5	Q95Q95	Q95Q95 caenorhabdi
35	32	64.0	2695	5	Q01438	Q01438 caenorhabdi
36	31	62.0	98	4	Q96EA3	Q96EA3 homo sapien
37	31	62.0	143	10	Q9XH06	Q9XH06 arabidopsis
38	31	62.0	166	4	Q9NXB1	Q9NXB1 homo sapien
39	31	62.0	206	16	Q8XGA2	Q8XGA2 ralstonia s
40	31	62.0	257	11	Q9CTD0	Q9CTD0 mus musculu
41	31	62.0	263	16	Q8ZQ47	Q8ZQ47 salmonella
42	31	62.0	263	16	Q8Z7P9	Q8Z7P9 salmonella
43	31	62.0	292	11	Q91V79	Q91V79 mus musculu
44	31	62.0	301	11	Q8CI63	Q8CI63 mus musculu
45	31	62.0	326	16	Q8YP22	Q8YP22 anabaena sp

## ALIGNMENTS

## RESULT 1

095146 ID 095146 PRELIMINARY; PRT; 58 AA.  
AC 095146;  
DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE LAGE-2ALT protein isoform.  
GN LAGE-2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Wang R.-F., Johnston S.L., Zeng G., Topalian S.L.,  
RA Schwartzentruber D.J., Rosenberg S.A.;  
RT "A breast and melanoma-shared tumor antigen: T Cell Responses to  
RT Antigenic Peptides Translated from Different Open Reading Frames.";  
RL J. Immunol. 161:3596-3606(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Lethe B.G.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99454989; PubMed=10523621;  
RA De Smet C., Lurquin C., Lethe B., Martelange V., Boon T.;  
RT "DNA methylation is the primary silencing mechanism for a set of germ  
RT line- and tumor-specific genes with a CpG-rich promoter.";  
RL Mol. Cell. Biol. 19:7327-7335(1999).  
DR EMBL; AF038567; AAD05203.1; -;  
DR EMBL; AJ275977; CAB76944.1; -;  
SQ SEQUENCE 58 AA; 6188 MW; ED12057564BC7EF2 CRC64;

Query Match 100.0%; Score 50; DB 4; Length 58;  
Best Local Similarity 100.0%; Pred. No. 0.0048;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



"Comparative Genome Sequencing for Discovery of Novel Polymorphisms in  
RT Bacillus anthracis.";  
RL Science 296:2028-2033(2002).  
DR EMBL: AF065404; AAD32364.1; -;  
DR EMBL: AF011190; AAM26041.1; -;  
KW Plasmid.  
SQ SEQUENCE 310 AA; 35923 MW; 4FEC85A695F4664C CRC64;

Query Match 68.0%; Score 34; DB 2; Length 310;  
Best Local Similarity 63.6%; Pred. No. 66;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 1 MLMQAQALAF 11  
|:::| |  
DB 263 MLMSKAYGFL 273

RESULT 6  
Q92DV2 PRELIMINARY; PRT; 388 AA.  
AC Q92DV2;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Hypothetical protein lin0710.  
GN LIN0710.  
OS Listeria innocua.  
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
OX NCBI\_TaxID=1642;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CLIP 11262 / Serovar 6a;  
RX PubMed=11679669;  
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,  
RA Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,  
RA Charbit A., Chetoui F., Couve E., de Barua A., Dehoux P.,  
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,  
RA Entian K.-D., Fshi H., Garcia-del Portillo F., Garrido P.,  
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,  
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,  
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nediari H.,  
RA Nordstiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,  
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,  
RA Vazquez-Boland J.-A., Voss H., Wehlend J., Cossart P.;  
RT "Comparative genomics of Listeria species.";  
RL Science 294:849-852(2001).  
DR EMBL: AL596166; CAC95942.1; -;  
DR ListList; LIN00710; -;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 388 AA; 44468 MW; F2D7DB541656F9D5 CRC64;

Query Match 68.0%; Score 34; DB 16; Length 388;  
Best Local Similarity 54.5%; Pred. No. 83;  
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MLMQAQALAF 11  
|:::| |  
DB 346 MLLAQQMSFL 356

RESULT 7  
Q94GH7 PRELIMINARY; PRT; 462 AA.  
AC Q94GH7;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Putative lipid acyl hydrolase.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Lillipsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzeae; Oryza.  
OX NCBI\_TaxID=4530;

[1]  
RN SEQUENCE FROM N.A.  
RP STRAIN-cv. Nipponbare;  
RA Buell C.R., Yuan Q., Ouyang S., Moffat K.S., Hill J.N., Gansberger K.,  
RA Brenner M., Burgess S., Hance M., Shvartsbeyn M., Tsltrin T.,  
RA Riggs F., Hsiao J., Zismann V., Blunt S., Pal G., Vanaken S.E.,  
RA Utterback T.R., Feldblyum T.V., Quackenbush J., Salzberg S.L.,  
RA White O., Fraser C.M.;  
RT "Oryza sativa chromosome 3 BAC OSJNB0093E13 genomic sequence.";  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AC091123; AAK72885.1; -;  
DR Gramene; Q94GH7; -;  
DR InterPro; IPR002641; Patatin.  
DR Pfam; PF01734; Patatin; 1.  
KW Hydrolase.  
SQ SEQUENCE 462 AA; 47466 MW; E07302B064F9B48D CRC64;

Query Match 68.0%; Score 34; DB 10; Length 462;  
Best Local Similarity 63.6%; Pred. No. 98;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MLMQAQALAF 11  
|:::| |  
DB 135 MFSADALAF 145

RESULT 8  
Q9LZC7 PRELIMINARY; PRT; 837 AA.  
AC Q9LZC7;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Hypothetical 94.3 kDa protein (Fragment).  
GN F8f6\_10.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Bevan M., Peters S.A., van Staveren M., Dirkse W., Stiekema W.,  
RA Bancroft I., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AL162873; CAB85500.1; -;  
DR InterPro; IPR002885; PPR.  
DR Pfam; PF01535; PPR; 12.  
DR TIGRfams; TIGR00756; PPR; 10.  
KW Hypothetical protein.  
FT NON\_TER 837 837  
SQ SEQUENCE 837 AA; 94342 MW; 20E286705EC8AD87 CRC64;

Query Match 68.0%; Score 34; DB 10; Length 837;  
Best Local Similarity 60.0%; Pred. No. 1.8e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLMQAQALAF 10  
|:::| |  
DB 345 MMAQDAVTF 354  
RESULT 9  
Q9FFN1 PRELIMINARY; PRT; 896 AA.  
AC Q9FFN1;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

Best Local Similarity 72.7%; Pred. No. 94;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MLMAQALAF 11  
Db 240 MLVQAHAFL 250

RESULT 11  
Q8S8E4 PRELIMINARY; PRT; 318 AA.

AC Q8S8E4;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Expressed protein.  
GN AT2G36000.  
OS Arabidopsis thaliana (Mouse-ear cross).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucosids II; Brassicales; Brassicaceae; Arabidopsi.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanKen S.E.,  
RA Barnstead M.E., Mason T.M., Bowman C.L., Ronning C.M., Benito M.-I.,  
RA Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,  
RA Fraser C.M., Venter J.C.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC007017; AAD21457.2; -  
DR InterPro; IPR003690; mTERF.  
DR Pfam; PF02536; mTERF; 1.  
DR SMART; SM00733; mTERF; 5.  
SQ SEQUENCE 318 AA; 36437 MW; 9596C45BF0740852 CRC64;

Query Match 66.0%; Score 33; DB 10; Length 318;  
Best Local Similarity 77.8%; Pred. No. 11e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 MAQALAF 11  
Db 1 MQEALSEL 9

RESULT 12  
Q8LF84 PRELIMINARY; PRT; 320 AA.

AC Q8LF84;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Hypothetical protein.  
OS Arabidopsis thaliana (Mouse-ear cross).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucosids II; Brassicales; Brassicaceae; Arabidopsi.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,  
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;  
RT "Full-length messenger RNA sequences greatly improve genome  
RT annotation";  
RL Genome Biol. 0:0-0(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,  
RA Feldmann K.;  
RT "Full-length cDNA from Arabidopsis thaliana";  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY084991; AAM61550.1; -  
DR InterPro; IPR003690; mTERF.

Query Match 66.0%; Score 33; DB 16; Length 272;  
Best Local Similarity 60.0%; Pred. No. 1.9e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLMAQALAF 10  
Db 345 MMAQDAVTF 354

RESULT 10  
Q8FD18 PRELIMINARY; PRT; 272 AA.

AC Q8FD18;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Shikimate 5-dehydrogenase (EC 1.1.1.25).  
GN AROE OR C4042.  
OS Escherichia coli O6.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=217992;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O6:H1 / CFT073 / ATCC 700928;  
RX MEDLINE=22388234; PubMed=12471157;  
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,  
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,  
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,  
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;  
RT "Extensive mosaic structure revealed by the complete genome sequence  
RT of uropathogenic Escherichia coli";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).  
DR EMBL; AE016767; AAN82480.1; -  
KW Oxidoreductase; Complete proteome.  
SQ SEQUENCE 272 AA; 29414 MW; EED84A94E9E15CAA CRC64;

Query Match 66.0%; Score 33; DB 16; Length 272;

DE AT5G03800/MED24\_10.  
OS Arabidopsis thaliana (Mouse-ear cross).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucosids II; Brassicales; Brassicaceae; Arabidopsi.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Columbia;  
RX MEDLINE=97471969; PubMed=9330910;  
RA Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M.,  
RA Miyaajima N., Tabata S.;  
RT "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence  
RT features of the 1.6 Mb regions covered by twenty physically assigned  
RT P1 clones";  
RL DNA Res. 4:215-230(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RA Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,  
RA Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,  
RA Hayashizaki Y., Shinozaki K.;  
RT "Arabidopsis thaliana full-length cDNA";  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB005235; BAB08606.1; -  
DR EMBL; AK117204; BAC41880.1; -  
DR InterPro; IPR000345; CytC\_heme\_bind.  
DR InterPro; IPR002885; PPR.  
DR Pfam; PF01535; PPR; 12.  
DR TIGRFAMs; TIGR00756; PPR; 10.  
DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 896 AA; 101155 MW; A4D9C4DEF09287B6 CRC64;

```

DR Pfam: PF02536; mTERF. 1.
DR SMART: SM00733; Mterf; 5.
KW Hypothetical protein.
SQ SEQUENCE 320 AA; 36649 MW; 507671B8ABE711A CRC64;

Query Match
Best Local Similarity 66.0%; Score 33; DB 10; Length 320;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 MAQEALAF 11
   1 | | | | | |
Db 1 MQQEALSF 9
   1 | | | | | |

RESULT 13
Q9SJ50 PRELIMINARY; PRT; 333 AA.
AC Q9SJ50;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE At2G36000/F11F19.9.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Gibson H.A., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Yamada K., Yu G., Yu S., Shinozaki K.,
RA Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis cDNA clones."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Meyers M.C., Banh J., Bowser L.,
RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis ORF clones."
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY039888; AAK63992.1; -.
DR EMBL: AY039871; AAK63975.1; -.
DR EMBL: AY077666; AAL76144.1; -.
DR InterPro: IPR003690; mTERF.
DR Pfam: PF02536; mTERF; 1.
DR SMART: SM00733; Mterf; 5.
SQ SEQUENCE 333 AA; 37952 MW; 072192A86556889A CRC64;

Query Match
Best Local Similarity 66.0%; Score 33; DB 10; Length 333;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 MAQEALAF 11
   1 | | | | | |
Db 1 MQQEALSF 9
   1 | | | | | |

RESULT 14
Q8WU54 PRELIMINARY; PRT; 338 AA.
AC Q8WU54;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)

```

```

DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RC TISSUE=Placenta;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC021238; AAH21238.1; -.
DR Pfam: PF04326; DUF467; 1.
KW Hypothetical protein.
SQ SEQUENCE 338 AA; 38102 MW; F321EDDA06C592FE CRC64;

Query Match
Best Local Similarity 66.0%; Score 33; DB 4; Length 338;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 AQEALAF 11
   1 | | | | | |
Db 130 SQEALAF 137
   1 | | | | | |

RESULT 15
Q9A7F1 PRELIMINARY; PRT; 565 AA.
AC Q9A7F1;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Ubiquinol oxidase subunit I.
GN CCL172.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus."
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
CC 1- COFACTOR: TWO HEME (A AND A3) GROUPS AND COPPER B (BY SIMILARITY).
CC 1- PATHWAY: TERMINAL STEP IN THE RESPIRATORY CHAIN.
CC 1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC 1- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.
DR EMBL: AE005851; AAK23748.1; -.
DR HSP: PI8401; 1FFT.
DR TIGR: CCL172; -.
DR InterPro: IPR000883; COX1.
DR Pfam: PF00115; COX1; 1.
DR PRINTS: PR01165; CYCOXIDASE1.
DR PROSITE: PS00077; COX1; 1.
KW Copper; Heme; Oxidoreductase; Respiratory chain; Transmembrane;
KW Complete proteome.
SQ SEQUENCE 665 AA; 74415 MW; 2BE3CB29E126608C CRC64;

Query Match
Best Local Similarity 66.0%; Score 33; DB 16; Length 665;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLMQEALAF 10
   1 | | | | | |
Db 86 MMRAQQAFAF 95
   1 | | | | | |

```

us-09-807-512-11.rspt

Tue Oct 7 15:52:20 2003

Search completed: October 7, 2003, 13:30:58  
Job time : 21.3797 secs





Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 4 AQEALAF 11  
:|||||  
Db 130 SQEALAF 137

RESULT 4  
US-10-156-761-13725  
; Sequence 13725, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HAYTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 13725  
; LENGTH: 204  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-13725

Query Match 64.0%; Score 32; DB 15; Length 204;  
Best Local Similarity 70.0%; Pred No. 86;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
Oy 2 LMAQEALAF 11  
:|||||  
Db 111 LVAQERFAFL 120

RESULT 5  
US-10-177-293-274  
; Sequence 274, Application US/10177293  
; Publication No. US20030124128A1  
; GENERAL INFORMATION:  
; APPLICANT: Lillie, James  
; APPLICANT: Glatt, Karen  
; APPLICANT: Zhao, Xumei  
; APPLICANT: Gannavarpu, Manjula  
; APPLICANT: Kamatkar, Shubhangi  
; APPLICANT: Mertens, Maureen  
; APPLICANT: Myer, Vic  
; APPLICANT: Wang, Youzhen  
; APPLICANT: Xu, Yongyao  
; APPLICANT: Hoersch, Sebastian  
; APPLICANT: Monahan, John  
; APPLICANT: Meyers, Rachel E.  
; APPLICANT: Bast Jr., Robert C.  
; APPLICANT: Hortobagyi, Gabriel N.  
; APPLICANT: Pusztai, Lajos  
; APPLICANT: Meric, Funda  
; APPLICANT: Sahin, Aysegul  
; APPLICANT: Mills, Gordon B.  
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT  
; FILE REFERENCE: MRI-038  
; CURRENT APPLICATION NUMBER: US/10/177,293  
; CURRENT FILING DATE: 2002-06-21  
; PRIOR APPLICATION NUMBER: US 60/299,887  
; PRIOR FILING DATE: 2001-06-21

APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10332  
; LENGTH: 272  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-09-815-242-10332

Query Match 66.0%; Score 33; DB 9; Length 272;  
Best Local Similarity 72.7%; Pred. No. 74;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
Oy 1 MLMAQEALAF 11  
:|||||  
Db 240 MLVAQAHAFL 250

RESULT 3  
US-09-949-192-45  
; Sequence 45, Application US/09949192  
; Patent No. US20020142292A1  
; GENERAL INFORMATION:  
; APPLICANT: Parham, Christi L.  
; APPLICANT: Gorman, Daniel L.  
; APPLICANT: Kurata, Hirokazu  
; APPLICANT: Arai, Naoko  
; APPLICANT: Sana, Theodore R.  
; APPLICANT: Mattson, Jeanline D.  
; APPLICANT: Murphy, Erin E.  
; APPLICANT: Savkoor, Chetan  
; APPLICANT: Grein, Jeffery  
; APPLICANT: Smith, Kathleen M.  
; APPLICANT: McClanahan, Terrill K.  
; TITLE OF INVENTION: MAMMALIAN GENES; RELATED REAGENTS AND METHODS  
; FILE REFERENCE: DX01169K  
; CURRENT APPLICATION NUMBER: US/09/949,192  
; CURRENT FILING DATE: 2001-09-07  
; PRIOR APPLICATION NUMBER: 60/231,267  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 45  
; LENGTH: 891  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-949-192-45  
Query Match 66.0%; Score 33; DB 10; Length 891;  
Best Local Similarity 87.5%; Pred. No. 2.8e+02;

```
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 274
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-293-274

Query Match      64.0%; Score 32; DB 15; Length 257;
Best Local Similarity 70.0%; Pred. No. 1.le+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1  MLMAQELALAF 10
        :||||| |
Db      12  LLMAQEHLEF 21

RESULT 6
US-10-155-613-1
; Sequence 1, Application US/10155613
; Publication No. US20020192787A1
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; LAL, Preeti
; Corley, Neil C.
; Guegler, Karl J.
; Yue, Henry
; TITLE OF INVENTION: HUMAN NUCLEIC ACID METHYLASES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/155,613
; FILING DATE: 22-May-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/575,205
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 09/082,310
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0520 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 463 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; IMMEDIATE SOURCE:
; LIBRARY: BRSTNOT07
; CLONE: 2124957
; SEQUENCE DESCRIPTION: SEQ ID NO: 1 :
US-10-155-613-1

Query Match      64.0%; Score 32; DB 14; Length 463;
Best Local Similarity 70.0%; Pred. No. 2.le+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1  MLMAQELALAF 10
        :||||| |
Db      12  LLMAQEHLEF 21

RESULT 7
US-10-314-881-9
; Sequence 9, Application US/10314881
; Publication No. US20030162279A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria Alexandra
; APPLICANT: Williamson, Mark
; APPLICANT: Tsia, Fong-Ying
; APPLICANT: Rudolph-Owen, Laura A.
; TITLE OF INVENTION: 22438, 23553, 25278, and 26212 No. US20030162279A1el
; FILE REFERENCE: 35800/208398(5800-79)
; CURRENT APPLICATION NUMBER: US/10/314,881
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US/09/773,426
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: US 09/495,823
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Pfam consensus sequence for human sulfatase
US-10-314-881-9

Query Match      64.0%; Score 32; DB 12; Length 552;
Best Local Similarity 70.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      2  LMAQELALPL 11
        :|:|||| |
Db      262  LLADEALPPL 271

RESULT 8
US-09-764-864-1207
; Sequence 1207, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1207
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (100)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
```

Tue Oct 7 15:52:18 2003

us-09-807-512-11.rapb

US-09-764-864-1207

Query Match 62.0%; Score 31; DB 10; Length 100;  
Best Local Similarity 60.0%; Pred. No. 63;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MLMAQEALAF 10  
| : : : |  
DB 17 MAMSQESLTF 26

RESULT 9

US-09-815-923-2  
; Sequence 2, Application US/09815923  
; Publication No. US20020137644A1  
; GENERAL INFORMATION:  
; APPLICANT: Gill, Sarjeet S.  
; APPLICANT: Ross, Linda S.  
; TITLE OF INVENTION: The Regents of the University of California  
; TITLE OF INVENTION: Use of Insect Cell Membrane Transporters as No. US20020137644A1el  
; FILE REFERENCE: Target Sites for Insecticides  
; FILE REFERENCE: 023070-093800US  
; CURRENT APPLICATION NUMBER: US/09/815,923  
; CURRENT FILING DATE: 2001-03-23  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 619  
; TYPE: PRT  
; ORGANISM: Manduca sexta  
; FEATURE:  
; OTHER INFORMATION: vesicular acetylcholine transporter  
US-09-815-923-2

Query Match 62.0%; Score 31; DB 10; Length 619;  
Best Local Similarity 54.5%; Pred. No. 4.7e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MLMAQEALAF 11  
: : : : |  
DB 290 LMMSNALAF 300

RESULT 10

US-10-148-806-35  
; Sequence 35, Application US/10148806  
; Publication No. US20030138933A1  
; GENERAL INFORMATION:  
; APPLICANT: Bai, Chang  
; APPLICANT: Metzger, Michael  
; APPLICANT: Liu, Xiaomei  
; TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NHL, A DNA  
; TITLE OF INVENTION: HELICASE  
; FILE REFERENCE: 20385P  
; CURRENT APPLICATION NUMBER: US/10/148,806  
; CURRENT FILING DATE: 2002-06-05  
; PRIOR APPLICATION NUMBER: US00/33065  
; PRIOR FILING DATE: 2000-12-09  
; PRIOR APPLICATION NUMBER: 60/169,970  
; PRIOR FILING DATE: 1999-12-09  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 35  
; LENGTH: 780  
; TYPE: PRT  
; ORGANISM: Dictyostelium discoideum  
US-10-148-806-35

Query Match 62.0%; Score 31; DB 12; Length 780;  
Best Local Similarity 54.5%; Pred. No. 6e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MLMAQEALAF 11

DB 345 MLSESPLAF 355  
| : : : |

RESULT 11

US-09-881-752A-178  
; Sequence 178, Application US/09881752A  
; Patent No. US20020115078A1  
; GENERAL INFORMATION:  
; APPLICANT: Kleanthous, Harold  
; APPLICANT: Al-Garawi, Amal  
; APPLICANT: Miller, Charles  
; APPLICANT: Tomb, Jean-Francois  
; APPLICANT: Oomen, Raymond P.  
; TITLE OF INVENTION: Identification of Polynucleotides  
; TITLE OF INVENTION: Encoding No. US20020115078A1el Helicobacter Polypeptides in t  
; TITLE OF INVENTION: Genome  
; FILE REFERENCE: 06132/041002  
; CURRENT APPLICATION NUMBER: US/09/881,752A  
; CURRENT FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 08/833,457  
; PRIOR FILING DATE: 1997-04-01  
; NUMBER OF SEQ ID NOS: 370  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 178  
; LENGTH: 136  
; TYPE: PRT  
; ORGANISM: Helicobacter pylori  
US-09-881-752A-178

Query Match 60.0%; Score 30; DB 10; Length 136;  
Best Local Similarity 54.5%; Pred. No. 1.4e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MLMAQEALAF 11  
| : : : |  
DB 12 MLMSREVSFV 22

RESULT 12

US-10-156-761-15002  
; Sequence 15002, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, YADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 15002  
; LENGTH: 288  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-15002

Query Match 60.0%; Score 30; DB 15; Length 268;  
Best Local Similarity 54.5%; Pred. No. 3e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MLMAQEALAF 11  
| : : : |  
DB 159 LLTAQDATAFV 169

; ORGANISM: Streptococcus pneumoniae  
US-09-855-400-2

Query Match 60.0%; Score 30; DB 10; Length 274;  
Best Local Similarity 75.0%; Pred. No. 3e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 AQEALAF 11  
II:IIII:  
Db 94 AQKALAFI 101

## RESULT 15

US-09-815-915-5  
; Sequence 5, Application US/09815915  
; Patent No. US20020025931A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Meyers, Rachel  
; TITLE OF INVENTION: 3714, 16742, 23546, AND 13887 NOVEL  
; TITLE OF INVENTION: PROTEIN KINASE MOLECULES AND USES THEREFOR  
; FILE REFERENCE: 38155-20006.00  
; CURRENT APPLICATION NUMBER: US/09/815,915  
; CURRENT FILING DATE: 2001-03-23  
; PRIOR APPLICATION NUMBER: US 60/191,846  
; PRIOR FILING DATE: 2000-03-24  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 341  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-815-915-5

Query Match 60.0%; Score 30; DB 9; Length 341;  
Best Local Similarity 63.6%; Pred. No. 3.9e+02;  
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MLMQAQALAF 11  
II:IIII:  
Db 145 MLQSSALAF 155

Search completed: October 7, 2003, 13:58:55  
Job time: 12.9747 secs

## RESULT 13

US-09-815-242-13286  
; Sequence 13286, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; TITLE OF INVENTION: Prokaryotes  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13286  
; LENGTH: 274  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-815-242-13286

Query Match 60.0%; Score 30; DB 9; Length 274;  
Best Local Similarity 75.0%; Pred. No. 3e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 AQEALAF 11  
II:IIII:  
Db 94 AQKALAFI 101

## RESULT 14

US-09-855-400-2  
; Sequence 2, Application US/09855400  
; Patent No. US20020115075A1  
; GENERAL INFORMATION:  
; APPLICANT: Burnham, Martin Karl Russel  
; APPLICANT: Biswas, Sanjoy  
; APPLICANT: Chalker, Alison Francis  
; APPLICANT: Ingraham, Karen Anne  
; APPLICANT: Traini, Christopher Michael  
; APPLICANT: Warren, Patrick Vernon  
; TITLE OF INVENTION: nade  
; FILE REFERENCE: GM10197  
; CURRENT APPLICATION NUMBER: US/09/855,400  
; CURRENT FILING DATE: 2001-05-15  
; PRIOR APPLICATION NUMBER: US/09/250,677  
; PRIOR FILING DATE: 1999-02-16  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 274  
; TYPE: PRT

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**

Tue Oct 7 15:52:17 2003

```
RESULT 14
US-09-250-677-2
; Sequence 2, Application US/09250677
; Patent No. 6251631
; GENERAL INFORMATION:
; APPLICANT: Burnham, Martin Karl Russel
; APPLICANT: Biswas, Sanjoy
; APPLICANT: Chalker, Allison Francis
; APPLICANT: Ingraham, Karen Anne
; APPLICANT: Traini, Christopher Michael
; APPLICANT: Warren, Patrick Vernon
; TITLE OF INVENTION: nade
; FILE REFERENCE: GM10197
; CURRENT APPLICATION NUMBER: US/09/250,677
; CURRENT FILING DATE: 1999-02-16
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-250-677-2

Query Match          60.0%; Score 30; DB 3; Length 274;
Best Local Similarity 75.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      4 AQEALAFI 11
DB      94 AQKALAFI 101

RESULT 15
US-09-252-991A-20828
; Sequence 20828, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20828
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20828

Query Match          60.0%; Score 30; DB 4; Length 307;
Best Local Similarity 66.7%; Pred. No. 2.1e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 MLMAQEALA 9
DB      170 VLLAQQA 178

Search completed: October 7, 2003, 13:34:05
Job time : 6.98734 secs
```



; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 23139  
; LENGTH: 268  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-23139

Query Match 62.0%; Score 31; DB 4; Length 268;  
Best Local Similarity 66.7%; Pred. No. 1.1e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 LMAQEALAF 10  
Db 88 LMAQEALIF 96

## RESULT 10

US-09-252-991A-23186  
; Sequence 23186, Application US/09252991A  
; Patent No. 6551795

; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 23186  
; LENGTH: 270  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-23186

Query Match 62.0%; Score 31; DB 4; Length 270;  
Best Local Similarity 66.7%; Pred. No. 1.2e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 LMAQEALAF 10  
Db 90 LMAQEALIF 98

## RESULT 11

US-09-252-991A-26258  
; Sequence 26258, Application US/09252991A  
; Patent No. 6551795

; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 26258  
; LENGTH: 503  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-26258

Query Match 62.0%; Score 31; DB 4; Length 503;  
Best Local Similarity 60.0%; Pred. No. 2.2e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 LMAQEALAF 11  
Db 244 LLAHALAVL 253

## RESULT 12

US-09-198-452A-128  
; Sequence 128, Application US/09198452A  
; Patent No. 6559294

; GENERAL INFORMATION:  
; APPLICANT: Griflais, R.  
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection  
; FILE REFERENCE: 9710-003-999  
; CURRENT APPLICATION NUMBER: US/09/198,452A  
; CURRENT FILING DATE: 1998-11-24  
; NUMBER OF SEQ ID NOS: 6849  
; SEQ ID NO 128  
; LENGTH: 636  
; TYPE: PRT  
; ORGANISM: Chlamydia pneumoniae  
US-09-198-452A-128

Query Match 62.0%; Score 31; DB 4; Length 636;  
Best Local Similarity 75.0%; Pred. No. 2.8e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 MAQEALAF 10  
Db 58 LAQQALAF 65

## RESULT 13

US-09-738-946-12  
; Sequence 12, Application US/09738946  
; Patent No. 6579701

; GENERAL INFORMATION:  
; APPLICANT: EXELIXIS, INC.  
; TITLE OF INVENTION: DROSOPHILA HOMOLOGUES OF GENES AND PROTEINS IMPLICATED IN CANCER  
; FILE REFERENCE: EX00-043C  
; CURRENT APPLICATION NUMBER: US/09/738,946  
; CURRENT FILING DATE: 2000-12-14  
; PRIOR APPLICATION NUMBER: 60/170,832  
; PRIOR FILING DATE: 1999-12-14  
; PRIOR APPLICATION NUMBER: 60/170,838  
; PRIOR FILING DATE: 1999-12-14  
; PRIOR APPLICATION NUMBER: 60/178,580  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: 60/185,879  
; PRIOR FILING DATE: 2000-02-29  
; PRIOR APPLICATION NUMBER: 60/185,880  
; PRIOR FILING DATE: 2000-02-29  
; PRIOR APPLICATION NUMBER: 60/186,150  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: 60/189,701  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 12  
; LENGTH: 1901  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-738-946-12

Query Match 62.0%; Score 31; DB 4; Length 1901;  
Best Local Similarity 72.7%; Pred. No. 8.7e+02;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MLMAQEALAF 11  
Db 1711 MLTQLALAF 1721

```
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 LMAQAEALAF 10
Db 12 LLMAQEHLEF 21

RESULT 6
US-09-575-205-1
; Sequence 1, Application US/09575205
; Patent No. 6436683
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Yue, Henry
; TITLE OF INVENTION: HUMAN NUCLEIC ACID METHYLASES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/575,205
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/082,310
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0520 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 463 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRSTNOT07
; CLONE: 2124957
; US-09-575-205-1

Query Match 64.0%; Score 32; DB 4; Length 463;
Best Local Similarity 70.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LMAQAEALAF 10
Db 12 LLMAQEHLEF 21

RESULT 7
US-09-773-426A-9
; Sequence 9, Application US/09773426A
; Patent No. 6534302
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria Alexandra
; APPLICANT: Williamson, Mark
; APPLICANT: Tsia, Fong-Ying

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LMAQAEALAF 10
Db 12 LLMAQEHLEF 21

RESULT 8
US-09-252-991A-23138
; Sequence 23138, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23138
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-23138

Query Match 62.0%; Score 31; DB 4; Length 268;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 LMAQAEALAF 10
Db 88 LMAQAEALIF 96

RESULT 9
US-09-252-991A-23139
; Sequence 23139, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
```

; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Win95  
; SOFTWARE: Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/905,223  
; FILING DATE:  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelsen, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 408:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 92 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PROTEIN  
; ORIGINAL SOURCE:  
; ORGANISM: Homo Sapiens  
; TISSUE TYPE: Brain  
; FEATURE:  
; NAME/KEY: sig-peptide  
; LOCATION: -40..-1  
; IDENTIFICATION METHOD: Von Heijne matrix  
; OTHER INFORMATION: score 7.9  
; OTHER INFORMATION: seq SLLLLFGQFPASS/QE  
US-08-905-223-408

Query Match 64.0%; Score 32; DB 3; Length 92;  
Best Local Similarity 70.0%; Pred. No. 24;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 MLMAQELAF 10  
Db 12 LLMAQEHLEF 21

RESULT 3  
US-09-252-991A-19768  
; Sequence 19768, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR FILING DATE: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 19768  
; LENGTH: 173  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-19768

Query Match 64.0%; Score 32; DB 4; Length 173;  
Best Local Similarity 70.0%; Pred. No. 47;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 MLMAQELAF 10  
Db 19 MLASQELRF 28

RESULT 4  
US-09-328-352-7879

; Sequence 7879, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 7879  
; LENGTH: 360  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-7879

Query Match 64.0%; Score 32; DB 4; Length 360;  
Best Local Similarity 60.0%; Pred. No. 99;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 2 LMAQELAF 11  
Db 14 LMAQRLSFL 23

RESULT 5  
US-09-082-310-1  
; Sequence 1, Application US/09082310  
; Patent No. 6096526  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Lal, Preeti  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Yue, Henry  
; TITLE OF INVENTION: HUMAN NUCLEIC ACID METHYLASES  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/082,310  
; FILING DATE: Herewith  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CERRONE, MICHAEL C.  
; REGISTRATION NUMBER: 39,132  
; REFERENCE/DOCKET NUMBER: PF-0520 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 855-0555  
; TELEFAX: (650) 845-4166  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 463 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: BRSTNOT07  
; CLONE: 2124957  
US-09-082-310-1

Query Match 64.0%; Score 32; DB 3; Length 463;  
Best Local Similarity 70.0%; Pred. No. 1.3e+02;

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model  
Run on: October 7, 2003, 13:21:07 ; Search time 5.98734 Seconds  
(without alignments)  
77.734 Million cell updates/sec

Title: US-09-807-512-11  
Perfect score: 50  
Sequence: 1 MLMAQEAALFL 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues  
Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database :
- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*
  - 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*
  - 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*
  - 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*
  - 5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*
  - 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	70.0	338	4	US-09-252-991A-28498
2	32	64.0	92	3	US-08-905-223-408
3	32	64.0	173	4	US-09-252-991A-19768
4	32	64.0	360	4	US-09-328-352-7879
5	32	64.0	463	3	US-09-082-310-1
6	32	64.0	463	4	US-09-575-205-1
7	32	64.0	552	4	US-09-773-426A-9
8	31	62.0	268	4	US-09-252-991A-23138
9	31	62.0	268	4	US-09-252-991A-23139
10	31	62.0	270	4	US-09-252-991A-23186
11	31	62.0	503	4	US-09-252-991A-26258
12	31	62.0	636	4	US-09-198-452A-128
13	31	62.0	1901	4	US-09-738-946-12
14	30	60.0	274	3	US-09-250-677-2
15	30	60.0	307	4	US-09-252-991A-20828
16	30	60.0	415	3	US-09-176-657-8
17	30	60.0	415	4	US-09-421-299-8
18	30	60.0	763	4	US-09-252-991A-30146
19	29	58.0	118	2	US-08-588-258B-38
20	29	58.0	118	5	PCT-US96-08295-38
21	29	58.0	119	4	US-09-328-352-8051
22	29	58.0	153	4	US-08-936-165A-507
23	29	58.0	225	4	US-09-198-452A-648
24	29	58.0	394	3	US-08-791-115B-27
25	29	58.0	394	4	US-09-140-749-17
26	29	58.0	403	3	US-08-791-115B-1
27	29	58.0	403	3	US-08-791-115B-23

28	29	58.0	403	3	US-08-791-115B-25	Sequence 25, Appl
29	29	58.0	403	4	US-09-140-749-2	Sequence 2, Appl
30	29	58.0	403	4	US-09-140-749-10	Sequence 10, Appl
31	29	58.0	403	4	US-09-140-749-49	Sequence 49, Appl
32	29	58.0	403	4	US-09-140-749-55	Sequence 55, Appl
33	29	58.0	403	4	US-09-140-749-57	Sequence 57, Appl
34	29	58.0	430	3	US-08-791-115B-7	Sequence 7, Appl
35	29	58.0	430	4	US-09-140-749-18	Sequence 18, Appl
36	29	58.0	542	4	US-08-311-731A-213	Sequence 213, App
37	29	58.0	559	4	US-09-140-749-15	Sequence 15, Appl
38	29	58.0	565	3	US-08-906-156A-12	Sequence 12, Appl
39	29	58.0	597	4	US-09-140-749-8	Sequence 8, Appl
40	29	58.0	645	3	US-08-791-115B-6	Sequence 6, Appl
41	29	58.0	670	4	US-09-252-991A-24930	Sequence 24930, A
42	29	58.0	729	4	US-09-107-532A-6946	Sequence 6946, Ap
43	29	58.0	742	3	US-08-791-115B-5	Sequence 5, Appl
44	29	58.0	797	1	US-08-698-551-18	Sequence 18, Appl
45	29	58.0	797	2	US-08-839-032A-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1  
US-09-252-991A-28498  
; Sequence 28498, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252.991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 28498  
; LENGTH: 338  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-28498

Query Match 70.0%; Score 35; DB 4; Length 338;  
Best Local Similarity 63.6%; Pred. No. 24;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MLMAQEAALFL 11  
Db 167 MLFAQDALVYL 177

RESULT 2  
US-08-905-223-408  
; Sequence 408, Application US/08905223  
; Patent No. 6222029  
; GENERAL INFORMATION:  
; APPLICANT: Edwards, Jean-Baptiste D.  
; APPLICANT: Duclert, Aymeric  
; APPLICANT: Lacroix, Bruno  
; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS  
; NUMBER OF SEQUENCES: 503  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson & Bear  
; STREET: 501 West Broadway  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92101-3505  
; COMPUTER READABLE FORM: Floppy Disk  
; MEDIUM TYPE:

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:04:57 ; Search time 180.747 seconds  
(without alignments)  
95.721 Million cell updates/sec

Title: US-09-807-512-2  
Perfect score: 579  
Sequence: 1 MLMQEAFLAFLMAQAMLA.....SWSAGSCPGMPLSPDQGRF 109

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues  
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*  
1: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.\*  
2: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.\*  
3: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.\*  
4: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.\*  
5: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.\*  
6: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.\*  
7: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.\*  
8: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.\*  
9: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.\*  
10: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.\*  
11: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.\*  
12: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.\*  
13: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.\*  
14: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.\*  
15: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.\*  
16: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.\*  
17: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.\*  
18: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.\*  
19: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.\*  
20: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.\*  
21: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.\*  
22: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.\*  
23: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.\*  
24: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	579	100.0	109	21 AAY70854	Human CTL-recognis
2	264	45.6	58	20 AAY05966	Human cancer antig
3	264	45.6	58	21 AAY70863	Human tumour antig
4	264	45.6	58	23 AAU84819	Human NYNSO1b cons
5	177.5	30.7	3541	23 AAU85130	Human melanoma spe
6	152	26.3	30	23 AAU85115	Human NYNSO1b segm
7	133	23.0	30	23 AAU85116	Human NYNSO1b segm
8	131	22.6	30	23 AAU85114	Human NYNSO1b segm
9	101.5	17.5	392	22 ABG06587	Novel human diagno

10	96.5	16.7	19938	24	ABP76678	Streptomyces virid
11	92.5	16.0	101	22	ABG10976	Novel human diagno
12	92	15.9	19	21	AAY70905	Synthetic peptide,
13	91.5	15.8	699	22	AAU68556	Human novel cytoki
14	91.5	15.8	710	23	AAO15655	Human RaIGDS-like
15	91.5	15.8	710	24	ABP58338	Human cell growth,
16	91.5	15.8	710	24	ABP57918	Human Ras-like pro
17	91.5	15.8	715	22	AAU68555	Human novel cytoki
18	91	15.7	1694	19	AAW40109	Human alpha-6(IV)
19	90	15.5	235	22	ABG18963	Novel human diagno
20	89	15.4	1466	22	AAE02537	Porcine alpha1(III
21	88.5	15.3	19938	24	ABP76679	Streptomyces virid
22	87	15.0	414	22	AAU32999	Novel human secret
23	85.5	14.8	635	22	ABG20755	Novel human diagno
24	84	14.5	19	21	AAY70906	Synthetic peptide,
25	83.5	14.4	562	24	ABR41327	Human DTHP intrac
26	83.5	14.4	1093	14	AAK41001	Human myotonic dys
27	83	14.3	158	22	AAU67487	Propionibacterium
28	83	14.3	307	22	ABB65163	Drosophila melanog
29	83	14.3	19938	24	ABP76681	Streptomyces virid
30	83	14.3	19938	24	ABP76682	Streptomyces virid
31	82.5	14.2	524	22	ABG14128	Novel human diagno
32	82	14.2	135	24	ABR48211	Human bladder canc
33	82	14.2	135	24	ABU56509	Lung cancer-associ
34	82	14.2	135	24	ABU56695	Lung cancer-associ
35	82	14.2	180	19	AAW69664	Human LAGE-1 clone
36	82	14.2	180	21	AAY70860	Human LAGE-1 splic
37	82	14.2	180	23	ABB78346	Amino acid sequenc
38	82	14.2	180	23	AAU84820	Human LAGE-1 conse
39	82	14.2	180	24	ABP74199	Human LAGE-1a prot
40	82	14.2	210	19	AAW69663	Human LAGE-1 clone
41	82	14.2	210	21	AAY70861	Human LAGE-1 unsp1
42	82	14.2	210	24	ABP74200	Human LAGE-1b prot
43	81.5	14.1	212	24	ABP75469	Human secretory po
44	81.5	14.1	266	22	ABG18528	Novel human diagno
45	81.5	14.1	717	23	ABG79520	Human tumour suppl

ALIGNMENTS

RESULT 1  
AAY70854  
ID AAY70854 standard; Protein; 109 AA.  
XX AC AAY70854;  
XX 31-JUL-2000 (first entry)  
XX Human CTL-recognised Antigen on MELanoma (CAMEL) protein.  
DE CAMEL: CTL-recognised Antigen on MELanoma; cytotoxic T lymphocyte; CTL;  
KW tumour-associated antigen; LAGE-1; NY-ESO-1; anticancer; melanoma;  
KW cancer; immunotherapy; immunogenic peptide; immune response.  
XX  
OS Homo sapiens.  
XX  
FH Key  
FT Peptide  
FT Location/Qualifiers  
FT 1..11  
FT /label= CTL\_epitope  
FT /note= "Immunogenic peptide with ability to elicit a  
FT CTL response"  
FT 2..10  
FT /label= Immunogenic\_peptide  
FT /note= "Specific for HLA-A3"  
FT 2..11  
FT /label= CTL\_epitope  
FT /note= "Immunogenic peptide with ability to elicit a  
FT CTL response"  
FT 10..18  
FT /label= CAMEL\_10  
FT /note= "Specific for HLA-A2"  
FT 16..25  
FT Peptide

```

FT      /label= CAMEL16
FT      /note= "Specific for HLA-A2"
FT      17..25
FT      /label= CAMEL17
FT      /note= "Specific for HLA-A2"
FT      51..59
FT      /label= Immunogenic_peptide
FT      /note= "Specific for HLA-A3101"
FT      101..109
FT      /label= Immunogenic_peptide
XX      WO200023584-A1.
XX      27-APR-2000.
XX      15-OCT-1999; 99WO-EP07832.
XX      16-OCT-1998; 98EP-0119583.
XX      (BOEH ) BOEHRINGER INGELHEIM INT GMBH.
XX      (UYHO-) UNIV HOSPITAL LEIDEN.
XX      Schrier PI, Aarnoudse CA, Heider K, Klade C;
XX      WPI; 2000-339685/29.
XX      N-PSDB; AAD00149.
XX      Tumor-associated antigen useful for cancer immunotherapy is encoded by
XX      the open reading frame of LAGE-1 (a tumor-specific antigen) CDNA -
XX      Claim 1; Page 55; 73pp; English.
XX      The present protein sequence is the human tumour-associated antigen CAMEL
XX      (Cytotoxic T lymphocytes (CTL)-recognised Antigen on Melanoma). CAMEL
XX      protein is encoded by the LAGE-1 gene, a tumour-specific antigen. It is
XX      different from the LAGE-1 protein, since it is translated from a
XX      different open reading frame (ORF-1). It shows strong homology with
XX      NY-ESO-1, a melanoma specific tumour antigen. The tumour-associated
XX      antigen displayed on melanoma cells is recognised by cytotoxic T
XX      lymphocytes. CAMEL is expressed in tumour cell lines, tumour tissues
XX      (e.g. breast and lung) and in restricted number of healthy tissues. This
XX      sequence has anticancer activity. CAMEL tumour antigen and immunogenic
XX      peptides derived from it are useful for cancer immunotherapy. They have
XX      the potential to induce an immune response, by eliciting a CTL response.
XX      The DNA molecule is used to construct recombinant or fusion proteins.
XX      Query Match 100.0%; Score 579; DB 21; Length 109;
XX      Best Local Similarity 100.0%; Pred. No. 3.2e-52;
XX      Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 MLMAQALAFILMAQAGMLAAQERRVPRAAEVPGAQGGQGRGEEAPRGVMAVPLLRM 60
DB      1 MLMAQALAFILMAQAGMLAAQERRVPRAAEVPGAQGGQGRGEEAPRGVMAVPLLRM 60
QY      61 EGAPAGGGRTAACFCTSRCLSRPWRKRSWSAGSCPGMPHLSPDQGRF 109
DB      61 EGAPAGGGRTAACFCTSRCLSRPWRKRSWSAGSCPGMPHLSPDQGRF 109
XX      RESULT 2
XX      AAY05966
XX      ID AAY05966 standard; Protein; 58 AA.
XX      AC AAY05966;
XX      16-AUG-1999 (first entry)
XX      Human cancer antigen NY ESO-1/CAG-3 ORF2 protein.
XX      NY ESO-1/CAG-3 gene; CAG-3 gene; cancer antigen; human; leukaemia;
XX      non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;
KW      metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;
KW      uterine cancer; breast cancer; prostate cancer; ovarian cancer;
KW      cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;
KW      liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;
KW      vaccine; ORF2.
XX      Homo sapiens.
XX      WO9918206-A2.
XX      15-APR-1999.
XX      21-SEP-1998; 98WO-US19609.
XX      08-OCT-1997; 97US-0061428.
XX      (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX      Rosenberg SA, Wang RF;
XX      WPI; 1999-277270/23.
XX      N-PSDB; AAX58599.
XX      Cancer antigen NY ESO1/CAG-3
XX      Claim 5; Fig 3A; 88pp; English.
XX      The present sequence represents the ORF2 protein encoded by
XX      open reading frame 2 of the human ESO-1/CAG-3 (or CAG-3) gene.
XX      CAG-3 is a new and potent tumour antigen capable of eliciting an
XX      antigen specific immune response by T cells. Cancer peptides
XX      comprising ORF2, ORF1 (see AAY05966), portions of these peptides and
XX      their variants (see AAY05967-87), are useful as cancer vaccines that
XX      protect the recipient from development of cancer. The invention
XX      provides: vectors and host cells (also useful as vaccines); a
XX      method of diagnosis of cancer or precancer; a transgenic animal;
XX      antisense oligonucleotides that inhibit expression of the cancer
XX      peptide or tumour antigen; antibodies reacting with the CAG-3
XX      cancer peptide, useful in diagnostic and detection assays; and
XX      methods for preventing or inhibiting cancer by administering a
XX      cancer peptide, with or without an HLA molecule. The cancer
XX      peptides form part of, or are derived from, cancers such as
XX      primary or metastatic melanoma, thymoma, lymphoma, sarcoma, lung
XX      cancer, liver cancer, leukaemia, uterine cancer, cervical cancer,
XX      bladder cancer, kidney cancer and adenocarcinomas such as breast,
XX      prostate, ovarian, pancreatic and thyroid cancers. Melanoma is
XX      treated by inducing cancer-specific T cells in vitro for subsequent
XX      return to a patient.
XX      Query Match 45.6%; Score 264; DB 20; Length 58;
XX      Best Local Similarity 88.7%; Pred. No. 5.5e-20;
XX      Matches 55; Conservative 2; Mismatches 1; Indels 4; Gaps 1;
QY      1 MLMAQALAFILMAQAGMLAAQERRVPRAAEVPGAQGGQGRGEEAPRGVMAVPLLRM 60
DB      1 MLMAQALAFILMAQAGMLAAQERRVPRAAEVPGAQGGQGRGEEAPRGVMAVPLLRM 60
QY      61 EG 62
DB      57 QG 58
XX      RESULT 3
XX      AAY70863
XX      ID AAY70863 standard; Protein; 58 AA.
XX      AC AAY70863;
XX      31-JUL-2000 (first entry)
XX      Human tumour antigen, NY-ESO-1 short variant protein.

```

XX NY-ESO-1; CAMEL; CTL-recognised Antigen on MELANOMA; human; cancer; CTL;  
 KW cytotoxic T lymphocyte; tumour-associated antigen; LAGE-1; anticancer;  
 KW melanoma; immunotherapy; immune response; variant.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200023584-A1.  
 XX  
 PD 27-APR-2000.  
 XX  
 PF 15-OCT-1999; 99WO-EP07832.  
 XX  
 PR 16-OCT-1998; 98EP-0119583.  
 XX  
 PA (BOEH) BOEHRINGER INGELHEIM INT GMBH.  
 XX (UYHO-) UNIV HOSPITAL LEIDEN.  
 XX  
 PI Schrier PI, Aarnoudse CA, Heider K, Klade C;  
 XX  
 DR WPI; 2000-339685/29.  
 DR N-PSDB; AAD00152.  
 XX  
 PT Tumor-associated antigen useful for cancer immunotherapy is encoded by  
 PT the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA -  
 XX  
 PS Example 3; Page 65; 73pp; English.  
 XX  
 CC The present sequence is the human NY-ESO-1 short variant protein, a  
 CC tumour antigen, identified by screening an esophagus carcinoma cDNA  
 CC library. This protein is derived from open reading frame (ORF)-2,  
 CC that differs from the CAMEL protein (Cytotoxic T lymphocytes (CTL)-  
 CC recognised Antigen on MELANOMA), a tumour-associated antigen, by only its  
 CC last 5 amino acids. It contains epitopes of tumour specific T-cells.  
 CC NY-ESO-1 is expressed in different tumour types, but not in healthy  
 CC tissues except in testis. The tumour-associated antigen displayed on  
 CC melanoma cells is recognised by cytotoxic T lymphocytes. This sequence  
 CC has anticancer activity. CAMEL tumour antigen and immunogenic peptides  
 CC derived from it are useful for cancer immunotherapy. They have the  
 CC potential to induce an immune response, by eliciting a CTL response.  
 CC The DNA molecule is used for the construction of recombinant or fusion  
 CC proteins.  
 XX  
 SQ Sequence 58 AA;  
 Query Match 45.6%; Score 264; DB 21; Length 58;  
 Best Local Similarity 88.7%; Pred. No. 5.5e-20;  
 Matches 55; Conservative 2; Mismatches 1; Indels 4; Gaps 1;  
 QY 1 MLMQAEALFLMAQAGMLAAQERRVPRAAEVPGAQGGQGGPRGREGAPRGVMAVPLLRM 60  
 Db 1 MLMQAEALFLMAQAGMLAAQERRVPRAAEVPGAQGGQGGPRGREGAPRGVMAVPLLRM 60  
 QY 61 EG 62 : :  
 Db 57 QG 58 : :  
 RESULT 4  
 AAU84819  
 ID AAU84819 standard; Protein; 58 AA.  
 XX  
 AC AAU84819;  
 XX  
 DT 08-MAY-2002 (first entry)  
 XX  
 DE Human NYSOLB consensus sequence.  
 XX  
 KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;  
 KW viral infection; human immunodeficiency virus; melanoma;  
 KW bacterial infection; Salmonella; Legionella; parasitic infection;  
 KW Trypanosoma; Toxoplasma; Giardia.  
 XX

OS Homo sapiens.  
 XX  
 PN WO200190197-A1.  
 XX  
 PD 29-NOV-2001.  
 XX  
 PF 25-MAY-2001; 2001WO-AU00622.  
 XX  
 PR 26-MAY-2000; 2000AU-0007761.  
 XX  
 PA (AUSU) UNIV AUSTRALIAN NAT.  
 XX  
 PI Thomson SA, Ramshaw IA;  
 XX  
 DR WPI; 2002-147575/19.  
 XX  
 PT New synthetic polypeptides having several different segments of at  
 PT least one parent polypeptide linked together differently compared to  
 PT the linkage in the parent polypeptide, for inducing immune response  
 PT against a pathogen or cancer -  
 XX  
 PS Example 3; Fig 27; 364pp; English.  
 XX  
 CC The invention relates to a new synthetic polypeptide (I) comprising  
 CC several different segments of at least one parent polypeptide linked  
 CC together in a different relationship relative to their linkage in the  
 CC parent polypeptide to impede, abrogate or otherwise alter at least one  
 CC function associated with the parent polypeptide and for inducing an  
 CC immune response against a pathogen or cancer. Also included are a  
 CC synthetic polynucleotide encoding and a computer system for  
 CC designing the synthetic polypeptides. The synthetic polypeptides and  
 CC polynucleotides are referred to as a Savine. The synthetic polypeptide is  
 CC useful for modulating immune responses preferably directed against a  
 CC pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix,  
 CC colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone  
 CC liver, oesophagus, brain, testicle, uterus), as potentiating agents.  
 CC Compositions comprising the polypeptide may be used in the treatment or  
 CC prophylaxis against viral (such as infections caused by HIV (human  
 CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis  
 CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial  
 CC (e.g., infections caused by Neisseria, Meningococcus, Haemophilus,  
 CC Salmonella, Streptococcus, Legionella and Mycobacterium or parasitic  
 CC Trypanosoma, Toxoplasma and Giardia) infections. The present  
 CC sequence is a consensus sequence for a parent protein used to design a  
 CC Savine of the invention.  
 XX  
 SQ Sequence 58 AA;  
 Query Match 45.6%; Score 264; DB 23; Length 58;  
 Best Local Similarity 88.7%; Pred. No. 5.5e-20;  
 Matches 55; Conservative 2; Mismatches 1; Indels 4; Gaps 1;  
 QY 1 MLMQAEALFLMAQAGMLAAQERRVPRAAEVPGAQGGQGGPRGREGAPRGVMAVPLLRM 60  
 Db 1 MLMQAEALFLMAQAGMLAAQERRVPRAAEVPGAQGGQGGPRGREGAPRGVMAVPLLRM 60  
 QY 61 EG 62 : :  
 Db 57 QG 58 : :  
 RESULT 5  
 AAU85130  
 ID AAU85130 standard; Protein; 3541 AA.  
 XX  
 AC AAU85130;  
 XX  
 DT 08-MAY-2002 (first entry)  
 XX  
 DE Human melanoma specific Savine.  
 XX  
 KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;

Tue Oct 7 15:52:28 2003

KW viral infection; human immunodeficiency virus; melanoma;
KW bacterial infection; Salmonella; Legionella; parasitic infection;
KW Trypanosoma; Toxoplasma; Giardia.
XX Homo sapiens.
OS Synthetic.
OS WO200190197-A1.
XX PN
XX XX
XX 29-NOV-2001.
XX PD
XX 25-MAY-2001; 2001WO-AU00622.
XX PF
XX 26-MAY-2000; 2000AU-0007761.
XX PR
XX (AUSU ) UNIV AUSTRALIAN NAT.
XX PA
XX Thomson SA, Ramshaw IA;
XX PI
XX WPI; 2002-147575/19.
XX DR N-PSDB; ABK36950.
XX DR
XX New synthetic polypeptides having several different segments of at
XX PT least one parent polypeptide linked together differently compared to
XX PT the linkage in the parent polypeptide, for inducing immune response
XX PT against a pathogen or cancer
XX PT
XX Example 3; Fig 27; 364pp; English.
XX PS
XX The invention relates to a new synthetic polypeptide (I) comprising
XX CC several different segments of at least one parent polypeptide linked
XX CC together in a different relationship relative to their linkage in the
XX CC parent polypeptide to impede, abrogate or otherwise alter at least one
XX CC function associated with the parent polypeptide and for inducing an
XX CC immune response against a pathogen or cancer. Also included are a
XX CC synthetic polynucleotide encoding and a computer system for
XX CC designing the synthetic polypeptides. The synthetic polypeptides and
XX CC polynucleotides are referred to as a savine. The synthetic polypeptide is
XX CC useful for modulating immune responses preferably directed against a
XX CC pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, bone
XX CC colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone
XX CC liver, oesophagus, brain, testicle, uterus), as potentiating agents.
XX CC Compositions comprising the polypeptide may be used in the treatment or
XX CC prophylaxis against viral (such as infections caused by HIV (human
XX CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis
XX CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial
XX CC (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,
XX CC Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic
XX CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,
XX CC Trypanosoma, Toxoplasma and Giardia) infections. The present
XX CC sequence is a savine protein of the invention.
XX XX
XX Sequence 3541 AA;
XX SQ
Query Match 30.7%; Score 177.5; DB 23; Length 3541;
Best Local Similarity 50.0%; Pred. NO. 3.3e-09;
Matches 43; Conservative 3; Mismatches 25; Indels 15; Gaps 2;
Qy 28 AAEVPGAGQGGPRGEEAPRGVMAVPLLRMEGAPGAGTACFCSTSRCLSRPW 87
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2951 SAEEVPGAGQGGPRGEEAPRGVMAA---RLQGGAPRGHGANAQDGRCPGARRPD 3007
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 88 KR-----SWSAGSCPGMPH 101
| | | | |
Db 3008 SRLILPGRCAGARASGRPGGAPRGPH 3033
| | | | |
RESULT 6
AAU85115
ID AAU85115 standard; Peptide: 30 AA.
XX
XX AC AAU85115;
XX AC

DT 08-MAY-2002 (first entry)
XX XX
XX Human NYNSO1b segment 2.
XX KW
XX Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;
XX KW viral infection; human immunodeficiency virus; melanoma;
XX KW bacterial infection; Salmonella; Legionella; parasitic infection;
XX KW Trypanosoma; Toxoplasma; Giardia.
XX XX
XX OS Homo sapiens.
XX XX
XX WO200190197-A1.
XX PN
XX 29-NOV-2001.
XX PD
XX 25-MAY-2001; 2001WO-AU00622.
XX PF
XX 26-MAY-2000; 2000AU-0007761.
XX PR
XX (AUSU ) UNIV AUSTRALIAN NAT.
XX PA
XX Thomson SA, Ramshaw IA;
XX XX
XX PI
XX WPI; 2002-147575/19.
XX DR N-PSDB; ABK36935.
XX DR
XX New synthetic polypeptides having several different segments of at
XX PT least one parent polypeptide linked together differently compared to
XX PT the linkage in the parent polypeptide, for inducing immune response
XX PT against a pathogen or cancer
XX PT
XX Example 3; Fig 27; 364pp; English.
XX PS
XX The invention relates to a new synthetic polypeptide (I) comprising
XX CC several different segments of at least one parent polypeptide linked
XX CC together in a different relationship relative to their linkage in the
XX CC parent polypeptide to impede, abrogate or otherwise alter at least one
XX CC function associated with the parent polypeptide and for inducing an
XX CC immune response against a pathogen or cancer. Also included are a
XX CC synthetic polynucleotide encoding and a computer system for
XX CC designing the synthetic polypeptides. The synthetic polypeptides and
XX CC polynucleotides are referred to as a savine. The synthetic polypeptide is
XX CC useful for modulating immune responses preferably directed against a
XX CC pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, bone
XX CC colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone
XX CC liver, oesophagus, brain, testicle, uterus), as potentiating agents.
XX CC Compositions comprising the polypeptide may be used in the treatment or
XX CC prophylaxis against viral (such as infections caused by HIV (human
XX CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis
XX CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial
XX CC (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,
XX CC Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic
XX CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,
XX CC Trypanosoma, Toxoplasma and Giardia) infections. The present
XX CC sequence is a peptide derived from a parent protein used to
XX CC construct a savine of the invention.
XX XX
XX Sequence 30 AA;
XX SQ
Query Match 26.3%; Score 152; DB 23; Length 30;
Best Local Similarity 100.0%; Pred. NO. 1e-08;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 14 QGAMLAQERRVPRAAEVPGAQGGQGRGR 43
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 QGAMLAQERRVPRAAEVPGAQGGQGRGR 30
|||||:|||||:|||||:|||||:|||||:|||||:
RESULT 7
AAU85116
ID AAU85116 standard; Peptide: 30 AA.
XX
XX AC AAU85116;
XX AC





ABP76678 standard; Protein; 19938 AA.  
XX ABP76678;  
AC  
XX 26-FEB-2003 (first entry)  
DT  
XX Streptomyces viridochromogenes Avi gene cluster polypeptide frame 2.  
DE  
XX Avilamycin; antibacterial; virucide; protozoacide; fungicide; infection;  
KW medicine; Staphylococcus aureus; biosynthetic gene cluster; enzyme.  
XX  
XX Streptomyces viridochromogenes.  
OS  
XX WO200268436-A1.  
PN  
XX 06-SEP-2002.  
PD  
XX 24-AUG-2001; 2001WO-EP09815.  
PF  
XX 25-FEB-2001; 2001DE-1009166.  
PR  
XX (COMB-) COMBINATURE BIOPHARM AG.  
PA  
XX Weitnauer G, Muehlenweg A, Trefzer A, Bechthold A;  
PI  
XX WPI: 2003-018650/01.  
DR N-PSDB; AB237515.  
XX  
XX New avilamycin derivatives, useful for treatment of infections, and  
PT nucleic acid encoding avilamycin synthesis enzymes -  
PT  
XX Example 1; Page 68-301; 319pp; German.  
PS  
XX The invention relates to avilamycin derivatives (I) with antibacterial,  
CC virucide, protozoacide and fungicide activity. (I) are useful for  
CC treatment of infections (bacterial, viral, protozoal or fungal), in human  
CC or veterinary medicine, particularly where caused by Staphylococcus  
CC aureus. (I) are more hydrophilic than known avilamycins. The present  
CC sequence is that of an avilamycin synthesis enzyme from the Streptomyces  
CC viridochromogenes Avilamycin A biosynthetic gene cluster  
CC (AB237515-AB237516).  
XX  
XX Sequence 19938 AA;  
SQ  
Query Match 16.7%; Score 96.5; DB 24; Length 19938;  
Best Local Similarity 36.9%; Pred. No. 4.6;  
Matches 38; Conservative 6; Mismatches 38; Indels 21; Gaps 6;  
QY 13 AOGAMLAQAQERRVP--RAAEVPGAQGOQ--GPRGEEAPRGVMAVPLLRMEGAPAGP 68  
Dd 13354 ARGAAAGRRPLRPPXRPAPHRPCASGRXDWGPRNAERAGRAPR---PHPRRGGRAPRRPG 13410  
QY 69 -----GRTAACFSCTSRCLSRPKNRWSAGS-CPGMP 100  
Dd 13411 PPRSPAPGYGRRRRRCAGAAAR---DRGARRAHSGSRCPWP 13450

ABG06587;  
XX 13-FEB-2002 (first entry)  
DT  
XX Novel human diagnostic protein #6578.  
DE  
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
XX Homo sapiens.  
OS  
XX WO200175067-A2.  
PN  
XX 11-OCT-2001.  
PD  
XX 30-MAR-2001; 2001WO-US08631.  
PF  
XX 31-MAR-2000; 2000US-0540217.  
PR 23-AUG-2000; 2000US-0649167.  
XX  
XX (HYSE-) HYSEQ INC.  
PA  
XX Drmanac RT, Liu C, Tang YT;  
PI  
XX WPI: 2001-639362/73.  
DR N-PSDB; AAS70774.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -  
PT  
XX Claim 20; SEQ ID NO 36946; 103pp; English.  
PS  
XX The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 392 AA;  
SQ  
Query Match 17.5%; Score 101.5; DB 22; Length 392;  
Best Local Similarity 38.8%; Pred. No. 0.024;  
Matches 33; Conservative 9; Mismatches 38; Indels 5; Gaps 3;  
QY 19 AAQERRVPRAAEVPGAQGOQPRGEEAPRGVMAVPLLRMEGAPAGGGRTAACFST 78  
Dd 58 AAQQRGRAREEAGAQAQEEPRGPQERRR--RLAEG-LRRQCGHPAGRGRGYYRAA 114  
QY 79 SRCLSRR--PWKRSWSAGSCPGMP 101  
Dd 115 GPARGGRTPEWQSGGARTRTPSVGH 139



us-09-807-512-2.rag

Tue Oct 7 15:52:28 2003

546 PGDPSSPTSSVSPGPPSPSRSDAPAGSPASP 579

asthma; coagulation disorder; haemophilia; sepsis; nephritis; inflammatory bowel disease; food supplement; immunogen.

Homo sapiens.

WO200175093-A1.

11-OCT-2001.

30-MAR-2001; 2001WO-US10484.

31-MAR-2000; 2000US-0540217.

23-AUG-2000; 2000US-0649167.

22-SEP-2000; 2000US-0668680.

23-OCT-2000; 2000US-0695618.

30-NOV-2000; 2000US-0728711.

14-MAR-2001; 2000US-0728711.

(HYSE-) HYSEQ INC.

Tang YT, Asundi V, Zhou P, Xue AJ, Ren F, Zhang J, Wang J, Xu C;

Yang Y, Zaho QA, Chen R, Wang D, Goodrich RW, Liu C, Drmanac RT;

WPI: 2001-626432/72.

N-PSDB; AAS59848.

New polypeptides and nucleic acids, useful for diagnosis, treatment of inflammatory, autoimmune, neurological, myeloid or lymphoid cell, bone degenerative disorders, cancer and promoting wound healing

Claim 20; Page 279-280; 336pp; English.

The invention relates to isolated human polypeptides (which may be cytokines) and the polynucleotides encoding them. The protein is useful for identifying a compound which binds to it (e.g. modulators, agonists and antagonists). The polynucleotides are useful as an array for mismatch detection. The proteins and nucleic acids are useful as nutritional sources or supplements. The protein exhibits activity relating to cytokine, cell proliferation, cell differentiation, antiinflammatory, stem cell growth factor activity, immune stimulating or immune suppressing and activin or inhibin related activities. The proteins (and antibodies raised against them) and nucleic acids are therefore useful in the diagnosis and treatment of diseases and disorders such as cancer, central and peripheral nervous system diseases and neuropathies, Alzheimer's, Parkinson's disease, Huntington's disease, cerebrovascular lateral sclerosis, spinal cord disorders, head trauma, cerebrovascular diseases, stroke, myeloid or lymphoid cell disorders, platelet disorders, thrombocytopaenia, stem cell disorders, aplastic anaemia, for regeneration of bone, cartilage, tendon, ligament and/or nerve tissue growth, and in tissue repair, healing of burns, incisions, ulcers, for treating osteoporosis, osteoarthritis, bone degenerative disorders, or periodontal disease, lung or liver fibrosis, reperfusion injury in various tissues, various immune deficiencies and disorders including severe combined immunodeficiency (SCID), bacterial or fungal infections, autoimmune disorders (e.g. multiple sclerosis, rheumatoid arthritis, diabetes mellitus, myasthenia gravis), allergic reactions and conditions, such as asthma or other respiratory problems, coagulation disorders, haemophilia, septic shock, sepsis, arthritis, nephritis and inflammatory bowel disease, viral infection and are useful in altering bodily characteristics. The present sequence represents a novel protein of the invention.

Sequence 699 AA;

Query Match 15.8%; Score 91.5; DB 22; Length 699;

Best Local Similarity 30.9%; Pred. No. 0.47;

Matches 29; Conservative 9; Mismatches 55; Indels 1; Gaps 1;

8 LAFLMAQAGMLAAQERRVPRAAEVPGAGQGGPRGEEAPRGVMAVPLLRMEGAPAG- 66

486 LAALHAQNLTEEQSYRLSVIEPPAAACSPSPRRIRISLTKRLSAKLAREKSSPSGS 545

67 PGRRTAACFCTSRCLSRPWRKRSWSAGSCPGMP 100

Db

RESULT 14

AAOI5655

AAOI5655 standard; Protein; 710 AA.

XX AAO15655;

XX 31-OCT-2002 (first entry)

XX Human RalGDS-like protein 3 (RGL3).

XX Human; RalGDS-like protein 3; RGL3 associated disorder;

XX guanine nucleotide exchange factor; RGL3; gene therapy.

XX Homo sapiens.

XX EP1229132-A2.

XX 07-AUG-2002.

XX 25-JAN-2002; 2002EP-0001159.

XX 30-JAN-2001; 2001WO-US00663.

XX 30-JAN-2001; 2001WO-US00664.

XX 30-JAN-2001; 2001WO-US00665.

XX 30-JAN-2001; 2001WO-US00666.

XX 30-JAN-2001; 2001WO-US00667.

XX 30-JAN-2001; 2001WO-US00668.

XX 30-JAN-2001; 2001WO-US00669.

XX 30-JAN-2001; 2001WO-US00670.

XX 23-MAY-2001; 2001US-0864761.

XX 28-SEP-2001; 2001US-326105P.

XX (AEOM-) AEOMICA INC.

XX Gu Y, Nguyen C;

XX WPI: 2002-620726/67.

XX N-PSDB; AAL44435.

XX Novel RalGDS-like protein 3, a guanine nucleotide exchange factor for

XX small GTPase Ral and downstream effector for both Rit and Ras and

XX nucleic acid encoding it for diagnosing, treating disorders associated

XX with RGL3

XX Claim 13; Fig 3; 60pp; English.

XX The invention comprises the amino acid and coding sequence of the human

XX RalGDS-like protein 3 (RGL3). RGL3 is a guanine nucleotide exchange

XX factor for the small GTPase Ral and a downstream effector for both Rit

XX and Ras. The RGL3 DNA and protein sequences of the invention are useful

XX for the diagnosis and treatment/prevention (e.g. gene therapy) of a

XX disorder associated with decreased or increased expression or activity of

XX human RGL3. The present amino acid sequence represents the human RGL3

XX protein.

Query Match 15.8%; Score 91.5; DB 23; Length 710;

Best Local Similarity 30.9%; Pred. No. 0.48;

Matches 29; Conservative 9; Mismatches 55; Indels 1; Gaps 1;

8 LAFLMAQAGMLAAQERRVPRAAEVPGAGQGGPRGEEAPRGVMAVPLLRMEGAPAG- 66

481 LAALHAQNLTEEQSYRLSVIEPPAAACSPSPRRIRISLTKRLSAKLAREKSSPSGS 540

67 PGRRTAACFCTSRCLSRPWRKRSWSAGSCPGMP 100

541 PGDPSSPTSSVSPGPPSPSRSDAPAGSPASP 574



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:19:12 ; Search time 61.3987 Seconds  
(without alignments)  
170.726 Million cell updates/sec

Title: US-09-807-512-2  
Perfect score: 579  
Sequence: 1 MLMAQEAFLMAQAMLA.....SWSAGSCFPMHLSPDQGRF 109  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues  
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_76: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	85	14.7	674	S13301	collagen alpha 1(X)
2	82.5	14.2	1106	JQ0405	hypothetical 119.5
3	82	14.2	315	E75382	hypothetical prote
4	81	14.0	1298	EDBE75	immediate-early pr
5	80	13.8	313	T28465	hypothetical prote
6	80	13.8	324	T28032	hypothetical prote
7	79.5	13.7	471	S01037	flavonol 3-O-gluc
8	78.5	13.6	471	S01037	flavonol 3-O-gluc
9	78.5	13.6	471	S01052	flavonol 3-O-gluc
10	78.5	13.6	676	B47222	Kallmann syndrome
11	78	13.5	1466	CGHU7L	collagen alpha 1(I
12	77.5	13.4	290	A32249	collagen - sea urc
13	77.5	13.4	1414	S23809	collagen alpha 2(I
14	77	13.3	363	T16831	hypothetical prote
15	77	13.3	645	D90782	probable tail fibe
16	77	13.3	645	H85642	probable tail fibe
17	77	13.3	917	S09646	collagen alpha 2(V
18	77	13.3	1018	CGHU2A	collagen alpha 2(V
19	77	13.3	1049	CGHU2S	collagen alpha 1(I
20	77	13.3	1366	CGHU2S	collagen alpha 2(I
21	76.5	13.2	302	T21257	hypothetical prote
22	76	13.1	221	S46234	Gbpip protein - Ch
23	76	13.1	975	S33121	homeotic protein C
24	76	13.1	1029	S21369	collagen alpha 2(V
25	76	13.1	1487	CGHU6C	collagen alpha 1(I
26	75.5	13.0	247	B42856	ubiquitin carrier
27	75.5	13.0	677	S23296	collagen alpha 2(I
28	75.5	13.0	2274	T30258	adenomatous polypo
29	75	13.0	171	A45626	merozoite surface

30	75	13.0	455	1	XUBHFG	flavonol 3-O-gluc
31	75	13.0	1532	2	A61262	collagen alpha 1(X
32	74.5	12.9	361	2	H90877	probable tail fibe
33	74.5	12.9	437	2	E90996	probable tail fibe
34	74.5	12.9	439	2	A85741	hypothetical prote
35	74.5	12.9	458	2	T31631	hypothetical prote
36	74	12.8	197	2	C49247	merozoite surface
37	74	12.8	1024	2	S18251	collagen alpha 1(X
38	73.5	12.7	276	2	T15172	hypothetical prote
39	73.5	12.7	289	2	T44462	DNA-binding protei
40	73.5	12.7	428	2	AH3536	dihydrolipoamide S
41	73.5	12.7	604	2	T36966	hypothetical prote
42	73.5	12.7	1464	2	S59856	collagen alpha 1(I
43	73	12.6	381	2	S16506	hypothetical prote
44	73	12.6	680	1	CGHU1D	collagen alpha 1(X
45	73	12.6	920	2	B34493	collagen alpha 1(I

ALIGNMENTS

RESULT 1  
S13301 collagen alpha 1(X) chain precursor - bovine  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 21-Nov-1993 #sequence\_revision 23-Feb-1996 #text\_change 13-Aug-1999  
C:Accession: S13301  
R:Thomas, J. T.; Kwan, A. P. L.; Grant, M. E.; Boot-Handford, R. P.  
Biochem. J. 273, 141-148, 1991  
A:Title: Isolation of cDNAs encoding the complete sequence of bovine type X collagen  
A:Reference number: S13301; MUID:91113131; PMID:1703407  
A:Accession: S13301  
A:Molecule type: mRNA  
A:Residues: 1-674 <THO>  
A:Cross-references: EMBL:X53556; NID:g263; PIDN:CAA37624.1; PID:g264  
C:Genetics:  
A:Gene: COL10A1  
C:Superfamily: collagen alpha 1(VIII) chain; complement Clq carboxyl-terminal homology  
C:Keywords: coiled coil; extracellular matrix; glycoprotein; homotrimer  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-674/Product: collagen alpha 1(X) chain #status predicted <MAT>  
F:547-673/Domain: complement Clq carboxyl-terminal homology <CIO>

Query Match	14.7%	Score 85;	DB 2;	Length 674;
Best Local Similarity	29.4%	Pred. No. 2.9;		
Matches	30;	Conservative	2;	Mismatches 38; Indels 32; Gaps 5;
QY	26	PRAAEVPGAQGGQGGPRGEAPRGVMAVPLLRMEGAPA--GPGGRTAACFSCTSRCLS	83	
DB	149	PAGISVPGKPGQPGTG-EPGPRG-----FPGEKGTSGVPLNGKGMGHCTPC-----	197	
QY	84	RRPWKRSW-----SAGSCFPMHLSPDQG	107	
DB	198	RPGERGLPGQPGTGPFGPGVGVKRGKENGGLFGQGLKGDQG	238	

RESULT 2  
JQ0405 hypothethical 119.5K protein (uvrA region) - Micrococcus luteus  
N:Alternate names: ORF 1 protein  
C:Species: Micrococcus luteus, Micrococcus lysodeikticus  
C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 04-Feb-2000  
C:Accession: JQ0405  
R:Shiota, S.; Nakayama, H.  
Mol. Gen. Genet. 217, 332-340, 1989  
A:Title: Micrococcus luteus homolog of the Escherichia coli uvrA gene: identification  
A:Reference number: S04781; MUID:89364717; PMID:2549377  
A:Accession: JQ0405  
A:Molecule type: DNA  
A:Residues: 1-1106 <SHI>  
A:Cross-references: EMBL:X15867  
A:Note: this reading frame extends between two stop codons and does not begin with a  
A:Note: the gene encoding this protein overlaps uvrA gene

C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;

Query Match 14.2%; Score 82.5; DB 2; Length 1106;  
Best Local Similarity 33.7%; Pred. No. 7.6; 52; Indels 15; Gaps 4;  
Matches 35; Conservative

QY 9 AFLMAQAMLAQAQERRVPRAAEVPGAGQGGPRGEEAPRGVMAVPL-----LRRM 60  
DB 540 AHARGHVLGAADGCAADRGAGQGGPRGEEAPRGVMAVPL-----LRRM 60  
QY 61 EGAPAGPGGRTAACTSCRLSRPWRKRSWSAGSCPGMPLSP 104  
DB 597 --GPAHPGHTDRLRAGRRRLPRRAVHRPAPAGQPP--PHRDP 636

RESULT 3  
hypothetical protein - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 28-Jul-2000  
A:Accession: E75382  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
M.; Shen, H.O.; Venter, J.C.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
Science 286, 1571-1577, 1999  
A:Reference number: A75250; MUID:20036896; PMID:10567266  
A:Accession: E75382  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-315 <WHI>  
A:Cross-references: GB:AE001998; GB:AE000513; NID:6459302; PIDN:AAF11113.1; PID:645931  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DR1541  
A:Map position: 1  
C:Superfamily: Deinococcus radiodurans hypothetical protein DR1541

Query Match 14.2%; Score 82; DB 2; Length 315;  
Best Local Similarity 32.8%; Pred. No. 2.8;  
Matches 38; Conservative 9; Mismatches 39; Indels 30; Gaps 7;

QY 10 FLMAQAMLAQAQERRVPRAAEVPG-----AQGGQGGPRGEEAPR--GVRMAVPL 56  
DB 186 FLGANGYWARHEQPLGPAEIPAGFIPAGVATGVARCDQDARGR---PTGFSFRAASPL 242  
QY 57 LRRMEGAPAGGGR-TAACTSCRLSRPWRKRSWSAGSCP-----GMPHLSPL 104  
DB 243 LLVYLSDRQPGQRYVAAATPGDHRKLTR-----WAAAPQPYGIWTNQMPYLNP 292

RESULT 4  
EBDE75  
Immediate-early protein IE175 - human herpesvirus 1  
C:Species: human herpesvirus 1  
C:Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 16-Jun-2000  
A:Accession: A23510  
R:McGeoch, D.J.; Dolan, A.; Donald, S.; Brauer, D.H.K.  
Nucleic Acids Res. 14, 1727-1745, 1986  
A:Title: Complete DNA sequence of the short repeat region in the genome of herpes simple  
A:Reference number: A23510; MUID:86148504; PMID:3005980  
A:Accession: A23510  
A:Molecule type: DNA  
A:Residues: 1-1298 <MCG>  
A:Cross-references: GB:X14112; GB:D00317; GB:D00374; GB:840593; NID:gl944536; PIDN:CAA32  
C:Comment: This protein acts at the transcriptional regulatory level and is required for  
C:Genetics:  
A:Gene: IE3  
A:Map position: short repeat region (IR-s)  
C:Superfamily: herpesvirus immediate-early protein IE175  
C:Keywords: DNA binding; early protein; transcription regulation

Query Match 14.0%; Score 81; DB 1; Length 1298;

Best Local Similarity 29.2%; Pred. No. 12;  
Matches 35; Conservative 11; Mismatches 46; Indels 28; Gaps 6;

QY 4 AQBALAFMAQAQAMLAQAQERRVPRAAEVPGAGQGGPRGEEAPRGVMAVPLLRM 60  
DB 710 AADALAAAAAASPAREGRKRSKSPARPPGGGPRPKTKKSCADAP--GSDARAPLPAPA 768  
QY 61 -----EGAPAGPGGRTAACTSCRSR--CLSRRP-----WKRWSAGSCPGMPL 101  
DB 769 PPSTPPGPEPAPAPAAAPRAAAQAARPPVAVSRRAEGDPDPLGGWRR-----QPPGSPH 823

RESULT 5  
T26465  
hypothetical protein Y11D7A.11 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jan-2000  
C:Accession: T26465  
R:Steward, C.  
submitted to the EMBL Data Library, October 1998  
A:Reference number: Z20218  
A:Accession: T26465  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-313 <WML>  
A:Cross-references: EMBL:AL032632; PIDN:CAA21586.1; GSPDB:GN00022; CESP:Y11D7A.11  
A:Experimental source: clone Y11D7A  
C:Genetics:  
A:Gene: CESP:Y11D7A.11  
A:Map position: 4 54/3; 88/1  
A:Introns: 26/2; 54/3; 88/1  
C:Superfamily: unassigned collagens

Query Match 13.8%; Score 80; DB 2; Length 313;  
Best Local Similarity 34.4%; Pred. No. 4.3;  
Matches 22; Conservative 3; Mismatches 27; Indels 12; Gaps 2;

QY 26 PRAAEVPGAGQGGPRGEEAPRGVMA-VPLLRMEGAPAGPG-----GR7AA 73  
DB 241 PGLRGAPGATGSGQGRNDGRPCNGKGPAGPGDGFDAAGPGGADGEPGAGGGLCAKGE 300  
QY 74 CFSC 77  
DB 301 CSHC 304

RESULT 6  
T28032  
hypothetical protein ZK836.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jan-2000  
C:Accession: T28032  
R:McLay, K.  
submitted to the EMBL Data Library, August 1996  
A:Reference number: Z20459  
A:Accession: T28032  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-324 <WIL>  
A:Cross-references: EMBL:Z78201; PIDN:CAB01588.1; GSPDB:GN00023; CESP:ZK836.1  
A:Experimental source: clone ZK836  
C:Genetics:  
A:Gene: CESP:ZK836.1  
A:Map position: 5  
A:Introns: 283/2  
C:Superfamily: unassigned collagens

Query Match 13.8%; Score 80; DB 2; Length 324;  
Best Local Similarity 37.1%; Pred. No. 4.4;  
Matches 23; Conservative 4; Mismatches 17; Indels 18; Gaps 3;

QY 31 VPGAQGGQGGPRGEEAPRGVMAVPLLRMEGAPAGPGGRTAACTSCRSR---CLSRRPW 87



Db 272 VPGYGSAGPDG-EEGPRG-----PSGLPGKDAEYCKPCPTRDDGNSHRAW 316  
QY 88 KR 89  
Db 317 RR 318

RESULT 7  
S08325  
flavonol 3-O-glucosyltransferase (EC 2.4.1.91) (allele BzMcC2) - maize  
N:Alternate names: UDPglucose flavonoid glucosyl-transferase  
C:Species: Zea mays (maize)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: S08325  
R:Furtek, D.; Schiefelbein, J.W.; Johnston, F.; Nelson Jr., O.E.  
Plant Mol. Biol. 11, 473-481, 1988  
A:Title: Sequence comparisons of three wild-type Bronze-1 alleles from Zea mays.  
A:Reference number: S08324  
A:Accession: S08325  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-471 <FUR>  
A:Cross-references: EMBL:X13501; NID:g22361; PIDN:CAA31856.1; PID:g295854  
C:Genetics:  
A:Gene: Bz1  
A:Introns: 175/1  
C:Superfamily: flavonol O3-glucosyltransferase  
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 13.7%; Score 79.5; DB 1; Length 471;  
Best Local Similarity 24.8%; Pred. No. 6.7;  
Matches 35; Conservative 14; Mismatches 59; Indels 33; Gaps 5;  
QY 1 MLMQAEALAFMA-QGAML-----AAQERRVPRAAEVPGAQGOQGRGEEAPRGYRMA- 53  
Db 28 LLSIARALAAAAAPSGATLSFLSTASSLAQLRKASASAGHGLPGNLRFEVDPGAPAAE 87  
QY 54 --VPLLRM-----EGAPAGPGGRTAACF-----SCTSRCLSRPWP 87  
Db 88 ETVPPVPRQOLFMEAAEAGGVKAWLEAARAAAGGARVTCVVGDAFWMPADAAASAGAPW 147  
QY 88 KRSWSAGSCPGMPLSPDGR 108  
Db 148 VPVWTAASCALLAHIRTDLSLR 168

RESULT 8  
S01037  
flavonol 3-O-glucosyltransferase (EC 2.4.1.91) (allele Bz-W22) - maize  
N:Alternate names: UDPglucose flavonoid glucosyltransferase  
C:Species: Zea mays (maize)  
C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 12-Nov-1999  
C:Accession: S01037; S08326  
R:Ralston, E.J.; English, J.J.; Dooner, H.K.  
Genetics 119, 185-197, 1988  
A:Title: Sequence of three bronze alleles of maize and correlation with the genetic fine  
A:Reference number: S01037; MUID:88284304; PMID:3396861  
A:Accession: S01037  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-471 <RAL>  
A:Cross-references: EMBL:X07937; NID:g22209; PIDN:CAA30760.1; PID:g22210  
R:Furtek, D.; Schiefelbein, J.W.; Johnston, F.; Nelson Jr., O.E.  
Plant Mol. Biol. 11, 473-481, 1988  
A:Title: Sequence comparisons of three wild-type Bronze-1 alleles from Zea mays.  
A:Reference number: S08324  
A:Accession: S08326  
A:Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-471 <FUR>  
A:Cross-references: EMBL:X13502; NID:g22505; PIDN:CAA31857.1; PID:g22506  
C:Genetics:  
A:Gene: Bz1

A:Map position: 9  
A:Introns: 175/1  
C:Superfamily: flavonol O3-glucosyltransferase  
C:Keywords: flavonoid biosynthesis; glycosyltransferase; hexosyltransferase  
Query Match 13.6%; Score 78.5; DB 2; Length 471;  
Best Local Similarity 24.6%; Pred. No. 8.3;  
Matches 34; Conservative 14; Mismatches 57; Indels 33; Gaps 5;  
QY 1 MLMQAEALAFMA-QGAML-----AAQERRVPRAAEVPGAQGOQGRGEEAPRGYRMA- 53  
Db 28 LLSIARALAAAAAPSGATLSFLSTASSLAQLRKASASAGHGLPGNLRFEVDPGAPAAE 87  
QY 54 --VPLLRM-----EGAPAGPGGRTAACF-----SCTSRCLSRPWP 87  
Db 88 ETVPPVPRQOLFMEAAEAGGVKAWLEAARAAAGGARVTCVVGDAFWMPADAAASAGAPW 147  
QY 88 KRSWSAGSCPGMPLSPD 105  
Db 148 VPVWTAASCALLAHIRTD 165

RESULT 9  
S01052  
flavonol 3-O-glucosyltransferase (EC 2.4.1.91) (allele Bz-McC) - maize  
N:Alternate names: UDPglucose flavonoid glucosyltransferase  
C:Species: Zea mays (maize)  
C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 18-Jun-1999  
C:Accession: S01052; S08324  
R:Ralston, E.J.; English, J.J.; Dooner, H.K.  
Genetics 119, 185-197, 1988  
A:Title: Sequence of three bronze alleles of maize and correlation with the genetic  
A:Reference number: S01037; MUID:88284304; PMID:3396861  
A:Accession: S01052  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-471 <RAL>  
A:Cross-references: EMBL:X07940; NID:g22204; PIDN:CAA30761.1; PID:g22205  
R:Furtek, D.; Schiefelbein, J.W.; Johnston, F.; Nelson Jr., O.E.  
Plant Mol. Biol. 11, 473-481, 1988  
A:Title: Sequence comparisons of three wild-type Bronze-1 alleles from Zea mays.  
A:Reference number: S08324  
A:Accession: S08324  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-471 <FUR>  
A:Cross-references: EMBL:X13500; NID:g22364; PIDN:CAA31855.1; PID:g1030071  
C:Genetics:  
A:Gene: Bz1  
A:Map position: 9  
A:Introns: 175/1  
C:Superfamily: flavonol O3-glucosyltransferase  
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 13.6%; Score 78.5; DB 2; Length 471;  
Best Local Similarity 24.6%; Pred. No. 8.3;  
Matches 34; Conservative 14; Mismatches 57; Indels 33; Gaps 5;  
QY 1 MLMQAEALAFMA-QGAML-----AAQERRVPRAAEVPGAQGOQGRGEEAPRGYRMA- 53  
Db 28 LLSIARALAAAAAPSGATLSFLSTASSLAQLRKASASAGHGLPGNLRFEVDPGAPAAE 87  
QY 54 --VPLLRM-----EGAPAGPGGRTAACF-----SCTSRCLSRPWP 87  
Db 88 ETVPPVPRQOLFMEAAEAGGVKAWLEAARAAAGGARVTCVVGDAFWMPADAAASAGAPW 147  
QY 88 KRSWSAGSCPGMPLSPD 105  
Db 148 VPVWTAASCALLAHIRTD 165

RESULT 10  
B47222

A:Accession: PE0011  
A:Molecule type: DNA  
A:Residues: 1-176 <EN>  
A:Cross-references: GB:M26939; NID:g180813; PIDN:AAA52040.1; PID:g180814  
R:Toman, P.D.; Ricca, G.A.; de Crombrughe, B.  
Nucleic Acids Res. 16, 7201, 1988  
A:Title: Nucleotide sequence of a cDNA coding for the amino-terminal region of human  
A:Reference number: S01726; MUID:88303360; PMID:3405773  
A:Accession: S01726  
A:Molecule type: mRNA  
A:Residues: 1-170 <OM>  
A:Cross-references: EMBL:X07240; NID:g30060; PIDN:CAA30229.1; PID:g30061  
A:Note: the authors translated the codon CAG for residue 154 as His  
R:Janeczko, R.A.; Ramirez, F.  
Nucleic Acids Res. 17, 6742, 1989  
A:Title: Nucleotide and amino acid sequences of the entire human alpha-1 (III) collagen  
A:Reference number: S04887; MUID:89386015; PMID:2780304  
A:Accession: S04887  
A:Molecule type: mRNA  
A:Residues: 149-163, G, 164-240, 'D', 242-471, 'D', 473-487, 'L', 489, 'S', 491-613, 'Y', 615-666  
A:Cross-references: EMBL:X15332; NID:g29545; PIDN:CAA33387.1; PID:g30045  
A:Note: the authors' translation of residues 905-932 is inconsistent with the nucleot  
R:Seyer, J.M.; Kang, A.H.  
Biochemistry 16, 1158-1164, 1977  
A:Title: Covalent structure of collagen: amino acid sequence of cyanogen bromide pept  
A:Reference number: A90399; MUID:77134724; PMID:557335  
A:Accession: A90399  
A:Molecule type: protein  
A:Residues: 'Y', 169-225, 229-232, 'P', 234-292, 'D', 294, 'S', 296-398 <SEY2>  
A:Experimental source: liver  
A:Note: sequence corrected by A94562; attachment of 2-O-alpha-D-glucosyl-O-beta-D-gal  
R:Seyer, J.M.  
submitted to the Atlas, December 1977  
A:Reference number: A94562  
A:Accession: A94562  
A:Molecule type: protein  
A:Residues: 'Y', 169-225, 229-277, 'A', 279-292, 'D', 294, 'S', 296-398 <SEY2>  
A:Experimental source: liver  
A:Note: author submitted corrections to A90399  
R:Milewicz, D.M.; Witz, A.M.; Smith, A.C.; Manchester, D.K.; Waldstein, G.; Byers, P  
Am. J. Hum. Genet. 53, 62-70, 1993  
A:Title: Parental somatic and germ-line mosaicism for a multilexon deletion with unusu  
ispring.  
A:Reference number: I51868; MUID:93304430; PMID:8317500  
A:Accession: I51868  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 186-194 <ML>  
A:Cross-references: GB:S62925; NID:g386425; PIDN:AAD13937.1; PID:g4261637  
R:Chiodo, A.A.; Sillence, D.O.; Cole, W.G.; Bateman, J.F.  
Biochem. J. 311, 939-943, 1995  
A:Title: Abnormal type III collagen produced by an exon-17-skipping mutation of the (C  
A:Reference number: S59511; MUID:96067614; PMID:7487954  
A:Accession: S59511  
A:Molecule type: mRNA  
A:Residues: 302-423 <CH1>  
A:Cross-references: GB:S79877; NID:g1195576; PIDN:AAB35615.1; PID:g1195577  
R:Seyer, J.M.; Kang, A.H.  
Biochemistry 17, 3404-3411, 1978  
A:Title: Covalent structure of collagen: amino acid sequence of five consecutive CNBR  
A:Reference number: A90414; MUID:79000343; PMID:687591  
A:Accession: A90414  
A:Molecule type: protein  
A:Residues: 399-675, 'N', 677-727 <SEY3>  
A:Experimental source: liver  
R:Lee, B.; Vitale, E.; Superti-Furga, A.; Steinmann, B.; Ramirez, F.  
J. Biol. Chem. 266, 5256-5259, 1991  
A:Title: G to T transversion at position +5 of a splice donor site causes skipping o  
A:Reference number: I55349; MUID:91161621; PMID:1672129  
A:Accession: I55349  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 537-605 <LEE>



C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001  
C:Accession: D90782  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: D90782  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-645 <HAY>  
A:Cross-references: GB:BA000007; PIDN:BA034651.1; PID:g13360688; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain RIMD 0509932  
C:Genetics:  
A:Gene: ECs1228

Query Match 13.3%; Score 77; DB 2; Length 645;  
Best Local Similarity 33.8%; Pred. No. 15;  
Matches 24; Conservative 13; Mismatches 18; Indels 16; Gaps 5;  
QY 3 MAQEALAFILM--AQGAMLAQE--RRVPRAAEVP-GAQGQGGPRGEEAPRGVRMAYPLLR 58  
Db 213 LAQESAGEILKRAEAATVSAEARRMAENARGPRGPQGETGPKG----- 262  
QY 59 RMEGAPAGPGG 69  
Db 263 --ETGPGVGPQG 271

Search completed: October 7, 2003, 13:32:32  
Job time : 64.3987 secs

254 ---LTGEPGKR 261

RESULT 13  
S23809  
collagen alpha 2(I) chain homolog - sea urchin (Strongylocentrotus purpuratus)  
C:Species: Strongylocentrotus purpuratus (purple urchin)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: S23809  
R:Exposito, J.Y.; d'Alessio, M.; Solursh, M.; Ramirez, F.  
J. Biol. Chem. 267, 15559-15562, 1992  
A:Title: Sea urchin collagen evolutionarily homologous to vertebrate pro-alpha-2(I) coll  
A:Reference number: S23809; MUID:92348411; PMID:1639795  
A:Accession: S23809  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1414 <EXP>  
A:Cross-references: EMBL:M92040; NID:g161435; PIDN:AAA30035.1; PID:g161436  
C:Superfamily: collagen alpha 2(I) chain; fibrillar collagen carboxyl-terminal homology  
C:Keywords: coiled coil; extracellular matrix; glycoprotein; trimer; triple helix  
F:1207-1414/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

Query Match 13.4%; Score 77.5; DB 1; Length 1414;  
Best Local Similarity 33.0%; Pred. No. 26;  
Matches 34; Conservative 4; Mismatches 42; Indels 23; Gaps 4;  
QY 13 AQGAMLAQERRVPRAAEVPAQGGQGGPRGEEA-----PRGVRMAVPL-LRRMEGAPA 65  
Db 406 AQGESGPLGRGTGTPAGPPGAQGESGERSGALGPAGPPGVGGERGMPGPMGSGAPG 465  
QY 66 GPG-----GRTAACFCTSRCLSRPWRKRSWAGSCPCMP 100  
Db 466 APGAGDKRLPGRSAGSKSAGESRPG-----PCMP 500

RESULT 14  
T16831  
hypothetical protein T07H6.3 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 18-Feb-2000  
C:Accession: T16831  
R:Geisel, C.  
submitted to the EMBL Data Library, April 1996  
A:Description: The sequence of C. elegans cosmid T07H6.  
A:Reference number: Z16586  
A:Accession: T16831  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-363 <GEI>  
A:Cross-references: EMBL:U53344; NID:g1255886; PID:g1255887; PIDN:AAA96223.1; GSPDB:GN00  
A:Experimental source: strain Bristol N2; clone T07H6  
C:Genetics:  
A:Gene: CESP:T07H6.3  
A:Map position: x  
A:Introns: 62/3  
C:Superfamily: unassigned collagens

Query Match 13.3%; Score 77; DB 2; Length 363;  
Best Local Similarity 30.1%; Pred. No. 9;  
Matches 25; Conservative 5; Mismatches 27; Indels 26; Gaps 3;  
QY 25 VPRAAEVPAQGGQGGPRGEEAPRGVRMAYPLLRMEGAPAGPGGRTAACFCTSRCLSR 84  
Db 284 VPGPAGAFAGPGDPGAPAGSP-----GAPNPGSG-----PQ 318  
QY 85 RPWRKRSWAGSCPCMPHLSPDOG 107  
Db 319 GPAGDNGGAGS-PCQPGANGDNG 340

RESULT 15  
D90782  
Probable tail fiber protein [imported] - Escherichia coli (strain O157:H7, substrain RIM

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:05:57 ; Search time 31.7342 Seconds  
(without alignments)  
161.527 Million cell updates/sec

Title: US-09-807-512-2  
Perfect score: 579  
Sequence: 1 MLMAQEAFLAFLAQAQMLAA.....SWSAGSCFMPHLSPDQGRF 109

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues  
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	85	14.7	674	1 CA1A_BOVIN	P23206 bos taurus
2	82	14.2	210	1 CTG2_HUMAN	O75638 homo sapien
3	81.5	14.1	339	1 CLR3_HUMAN	O9nyq7 homo sapien
4	81	14.0	339	1 SRR1_HUMAN	O9uh36 homo sapien
5	81	14.0	1298	1 ICP4_HSV1	P08392 herpes simp
6	80	13.8	492	1 DYJ2_HUMAN	O43237 homo sapien
7	79.5	13.7	471	1 UF02_MAIZE	P16165 zea mays (m
8	78.5	13.6	471	1 UF01_MAIZE	P16166 zea mays (m
9	78.5	13.6	471	1 UF03_MAIZE	P16167 zea mays (m
10	78.5	13.6	508	1 OT01_ONCKE	P83371 oncorhynch
11	78.5	13.6	676	1 KALM_CHICK	P33005 gallus gall
12	78	13.5	1466	1 CA13_HUMAN	P02461 homo sapien
13	77	13.3	180	1 CTG1_HUMAN	P78358 homo sapien
14	77	13.3	1019	1 CA26_HUMAN	P12110 homo sapien
15	77	13.3	1049	1 CA13_BOVIN	P04258 bos taurus
16	77	13.3	1366	1 CA21_HUMAN	P08123 homo sapien
17	76	13.1	425	1 Y450_HUMAN	O75038 homo sapien
18	76	13.1	747	1 CA12_BOVIN	P02459 bos taurus
19	76	13.1	975	1 CUT1_CANFA	P39881 canis famil
20	76	13.1	1029	1 CA26_MOUSE	O02788 mus musculu
21	76	13.1	1418	1 CA12_HUMAN	P02458 homo sapien
22	76	13.1	1804	1 CA1B_MOUSE	O61245 mus musculu
23	75.5	13.0	222	1 UBCE_HUMAN	Q16763 homo sapien
24	75.5	13.0	345	1 FX11_HUMAN	Q12952 homo sapien
25	75.5	13.0	2003	1 NT04_HUMAN	O99466 homo sapien
26	75	13.0	455	1 UF0G_HORVU	P14726 hordeum vul
27	75	13.0	1356	1 CA21_ONCMY	O92484 oncorhynch
28	74	12.8	911	1 CA1B_BOVIN	O28083 bos taurus
29	73.5	12.7	912	1 ICA5_RABIT	O28730 oryctolagus
30	73.5	12.7	1464	1 CA13_MOUSE	P08121 mus musculu
31	73.5	12.7	1533	1 PUM_DROME	P25822 drosophila
32	73	12.6	680	1 CA1A_HUMAN	O03692 homo sapien
33	73	12.6	1239	1 NME3_MOUSE	Q01098 mus musculu

ALIGNMENTS

RESULT 1			
CA1A_BOVIN	STANDARD;	PRT;	674 AA.
ID	CA1A_BOVIN		
AC	P23206;		
DT	01-NOV-1991 (Rel. 20, Created)		
DT	01-NOV-1991 (Rel. 20, Last sequence update)		
DT	15-JUL-1999 (Rel. 38, Last annotation update)		
DE	Collagen alpha 1(X) chain precursor.		
GN	COL10A1.		
OS	Bos taurus (Bovine).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;		
OC	Bovidae; Bovinae; Bos.		
OX	NCBI_TaxID=9913;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Cartilage;		
RX	MEDLINE=91113131; PubMed=1703407;		
RA	Thomas J.T., Kwan A.P.L., Grant M.E., Boot-Handford R.P.;		
RT	"Isolation of cDNAs encoding the complete sequence of bovine type X collagen. Evidence for the condensed nature of mammalian type X		
RT	collagen genes."		
RL	Biochem. J. 273:141-148(1991).		
CC	-!- FUNCTION: TYPE X COLLAGEN IS A PRODUCT OF HYPERTHROPHIC		
CC	CHONDROCYTES AND HAS BEEN LOCALIZED TO PRESUMPTIVE		
CC	MINERALIZATION ZONES OF HYALINE CARTILAGE.		
CC	-!- SUBUNIT: Homotrimer.		
CC	-!- PTM: Prolines at the third position of the tripeptide repeating		
CC	unit (G-X-Y) are hydroxylated in some or all of the chains.		
CC	-!- SIMILARITY: STRONG, TO ALPHA 1 AND 2 TYPE VIII COLLAGENS.		
CC	-!- SIMILARITY: Contains 1 C1Q domain.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>		
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).		
CC	-----		
CC	EMBL; X53556; CAA37624.1; -.		
DR	PIR; S13301; S13301.		
DR	InterPro; IPR001073; Clq.		
DR	InterPro; IPR000087; Collagen.		
DR	Pfam; PF00386; Clq; 1.		
DR	Pfam; PF01391; Collagen; 9.		
DR	PRINTS; PR00007; COMPLEMENTC1Q.		
DR	PRODom; PD000007; Clq_helix; 2.		
DR	SMART; SM00110; C1Q; 1.		
DR	PROSITE; PS01113; Clq; 1.		
KW	Extracellular matrix; Connective tissue; Repeat; Hydroxylation;		
KW	Cartilage; Collagen; Signal; Glycoprotein.		
FT	SIGNAL 1 18		
FT	CHAIN 19 674		
FT	DOMAIN 19 56		
FT	DOMAIN 57 519		
FT	COLLAGEN ALPHA 1(X) CHAIN.		
FT	NONHELICAL REGION (NC2).		
FT	TRIPLE-HELICAL REGION.		

P02457 gallus gall  
Q14055 homo sapien  
P02452 homo sapien  
P12111 homo sapien  
Q91714 mycobacteri  
P18832 caenorhabdi  
P31870 hepatitis b  
Q60474 cavia porce  
O75949 homo sapien  
P28481 mus musculu  
P12107 homo sapien  
Q9C009 homo sapien

34 73 12.6 1453 1 CA11\_CHICK  
35 72.5 12.5 689 1 CA29\_HUMAN  
36 72.5 12.5 1464 1 CA11\_HUMAN  
37 72.5 12.5 3176 1 CA36\_HUMAN  
38 72 12.4 181 1 Y004\_MYCPA  
39 72 12.4 316 1 CC07\_CAEEL  
40 72 12.4 842 1 DPOL\_HPBVM  
41 71.5 12.3 450 1 A2AA\_CAVPO  
42 71.5 12.3 473 1 TED\_HUMAN  
43 71.5 12.3 1459 1 CA12\_MOUSE  
44 71.5 12.3 1806 1 CA1B\_HUMAN  
45 71 12.3 403 1 FXQ1\_HUMAN

FT DOMAIN 520 674 NONHELICAL REGION (NC1).  
FT DOMAIN 539 674 C10.  
FT DISULFID 194 197 BY SIMILARITY.  
FT MOD\_RES 460 460 HYDROXYLATION (BY SIMILARITY).  
FT MOD\_RES 463 463 HYDROXYLATION (BY SIMILARITY).  
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 674 AA; 6546 MW; CD4CA73A03E004CA CRC64;  
  
Query Match 14.7%; Score 85; DB 1; Length 674;  
Best Local Similarity 29.4%; Pred. No. 2.7;  
Matches 30; Conservative 2; Mismatches 38; Indels 32; Gaps 5;  
  
QY 26 PRAAEVCAQGGQPGREAPRGVMAVPLLRMEGAPGAGTAACTSCRLSRLS 83  
DB 149 PAGISVPGKPGQPGTG-EPGPRG-----FFGEGKGTSGVGLNGKGMGHCTPC----- 197  
  
QY 84 RPPKRSW-----SAGSCPGMPLSPDQG 107  
DB 198 -RPGERGLPGQGTGPPGPGVCKRGENGLPQGPGLKGDQ 238  
  
RESULT 2  
CTG2\_HUMAN STANDARD; PRT; 210 AA.  
AC 075638; 075637;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cancer/testis antigen 2 (LAGE-1 protein).  
GN CTAG2 OR LAGE1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
ON NCBI\_TaxId=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORMS LAGE-1A AND LAGE-1B).  
RC Tissue=Melanoma;  
RX MEDLINE=98289662; PubMed=9626360;  
RA Lethe B., Lucas S., Michaux L., de Smet C., Godelaine D., Serrano A.,  
de Plaen E., Boon T.;  
RT "LAGE-1, a new gene with tumor specificity";  
RL Int. J. Cancer 76:903-908(1998).  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event-Alternative splicing; Named isoforms-2;  
CC Name=LAGE-1B;  
CC IsoId=075638-1; Sequence=Displayed;  
CC Name=LAGE-1A;  
CC IsoId=075638-2; Sequence=VSP\_004301;  
CC -1- TISSUE SPECIFICITY: TESTIS AND VERY LOW LEVEL IN PLACENTA AND IN  
CC SOME UTERUS SAMPLES. OBSERVED IN 25-50% OF TUMOR SAMPLES OF  
CC MELANOMAS, NON-SMALL-CELL LUNG CARCINOMAS, BLADDER, PROSTATE AND  
CC HEAD AND NECK CANCERS.  
CC -1- DOMAIN: A TRANSMEMBRANE DOMAIN IS PRESENT IN ISOFORM LAGE-1A.  
CC -1- SIMILARITY: BELONGS TO THE CTAG FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; AJ223093; CAAL1117.1; -  
CC DR EMBL; AJ223093; CAAL1116.1; -  
CC DR EMBL; AJ223040; CAAL1043.1; -  
CC DR EMBL; AJ223041; CAAL1044.1; -  
CC DR Genew; HNCN:2492; CTAG2.  
CC MIM; 300396; -  
CC Polymorphism; 5 79  
CC DOMAIN 183 188 GLY-RICH.  
FT FT POLY-PRO.  
FT VARSPIC 135 210 MSVMDQDREGAGRMVYVWGLSGSPGQKDLRTPKHV

FT FT SEQRPTGPPPPPEGAQDGCRCGVAFNMFSAPIH -> IR  
FT LTAADHROLOLSISSCLOLSLMLWITQCFLPVLQAQPSG  
FT QRR (in isoform LAGE-1A).  
FT /FTid=VSP\_004301.  
FT R -> Q.  
FT /FTid=VAR\_007855.  
FT E -> Q.  
FT /FTid=VAR\_007856.  
FT W -> R.  
FT /FTid=VAR\_007857.  
SQ SEQUENCE 210 AA; 21119 MW; 8BE0E00AE55E8BE CRC64;  
  
Query Match 14.2%; Score 82; DB 1; Length 210;  
Best Local Similarity 37.7%; Pred. No. 1.8;  
Matches 23; Conservative 3; Mismatches 19; Indels 16; Gaps 2;  
  
QY 33 GAQGGQGGPRG-----REEAPRGVMAVPLLRMEGAPGAGTAACTSCRLSRLSPWK 88  
DB 38 GATGGRGPRGAGARASGRG-----GAPRGPHGGAASQDGRCPGARRPDS 85  
  
QY 89 R 89  
DB 86 R 86  
  
RESULT 3  
CLR3\_HUMAN STANDARD; PRT; 3312 AA.  
AC Q9NYQ7; 075092;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cadherin EGF LAG seven-pass G-type receptor 3 precursor (Flamingo 2)  
DE homolog 1 (hFm1) (Multiple epidermal growth factor-like domains 2)  
DE (Epidermal growth factor-like 1).  
DE CELSR3 OR CDHF11 OR FMI1 OR EGFL1 OR MEGF2.  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
ON NCBI\_TaxId=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20202599; PubMed=10716726;  
RA Wu Q., Maniatis T.;  
RT "Large exons encoding multiple ectodomains are a characteristic  
RT feature of protocadherin genes";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:3124-3129(2000).  
RN [2]  
RP SEQUENCE OF 1954-3312 FROM N.A.  
RC Tissue=Brain;  
RX MEDLINE=98360089; PubMed=9693030;  
RA Nakayama M., Nakajima D., Nagase T., Nomura N., Seki N., Ohara O.;  
RT "Identification of high-molecular-weight proteins with multiple  
RT EGF-like motifs by motif-trap screening.";  
RL Genomics 51:27-34(1998).  
CC -1- FUNCTION: Receptor that may have an important role in cell/cell  
CC signaling during nervous system formation.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.  
CC -1- SIMILARITY: Contains 9 cadherin domains.  
CC -1- SIMILARITY: Contains 8 EGF-like domains.  
CC -1- SIMILARITY: Contains 2 laminin G-like domains.  
CC -1- SIMILARITY: Contains 1 laminin EGF-like domain.  
CC -1- SIMILARITY: Contains 1 GPS domain.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC

DR EMBL; AF231023; AAF61929.1; -.  
DR EMBL; AB011536; BAA32464.1; -.  
DR HSP; P00740; 1EDM.  
DR Genew; HGNC:3230; CELSR3.  
DR MIM; 604264; -.  
DR GO; GO:0005198; F:structural molecule activity; NAS.  
DR InterPro; IPR000152; ASX\_hydroxyl.  
DR InterPro; IPR002126; Cadherin.  
DR InterPro; IPR000742; EGF 2.  
DR InterPro; IPR001881; EGF Ca.  
DR InterPro; IPR006209; EGF like.  
DR InterPro; IPR000832; GPCR\_secretin.  
DR InterPro; IPR001879; hormn\_receptor.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR002049; Laminin\_EGF.  
DR InterPro; IPR001791; Laminin-G.  
DR InterPro; IPR000203; PKD\_cys\_rich.  
DR Pfam; PF00002; 7tm\_2; 1.  
DR Pfam; PF00028; cadherin; 9.  
DR Pfam; PF00008; EGF; 6.  
DR Pfam; PF01825; GPS; 1.  
DR Pfam; PF02793; HRM; 1.  
DR Pfam; PF00054; laminin.G; 2.  
DR PRINTS; PR00205; CADHERIN.  
DR PRINTS; PR00011; EGF\_LAMININ.  
DR PRINTS; PR00249; GPCRSECRETIN.  
DR SMART; SM00112; CA; 9.  
DR SMART; SM00181; EGF; 6.  
DR SMART; SM00303; GPS; 1.  
DR SMART; SM00008; HormR; 1.  
DR SMART; SM00282; LamG; 2.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 1.  
DR PROSITE; PS00232; CADHERIN.1; 7.  
DR PROSITE; PS0268; CADHERIN.2; 8.  
DR PROSITE; PS00022; EGF\_1; 6.  
DR PROSITE; PS01186; EGF\_2; 4.  
DR PROSITE; PS00221; GPS; 1.  
DR PROSITE; PS00649; G\_PROTEIN\_RECEP\_F2\_1; FALSE\_NEG.  
DR PROSITE; PS00650; G\_PROTEIN\_RECEP\_F2\_2; FALSE\_NEG.  
DR PROSITE; PS00227; G\_PROTEIN\_RECEP\_F2\_3; 1.  
DR PROSITE; PS00261; G\_PROTEIN\_RECEP\_F2\_4; 1.  
DR PROSITE; PS00025; LAM\_G\_DOMAIN; 2.  
DR PROSITE; PS01248; LAMININ\_TYPE\_EGF; 1.  
KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
KW EGF-like domain; Calcium-binding; Laminin EGF-like domain; Repeat;  
KW Developmental protein; Hydroxylation; Signal.  
FT SIGNAL 1 32 POTENTIAL.  
FT CHAIN 33 3312 CADHERIN EGF LAG SEVEN-PASS G-TYPE  
FT RECEPTOR 3.  
FT DOMAIN 33 2540 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 2541 2561 1 (POTENTIAL).  
FT DOMAIN 2562 2572 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 2573 2593 2 (POTENTIAL).  
FT DOMAIN 2594 2601 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 2602 2622 3 (POTENTIAL).  
FT DOMAIN 2623 2643 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 2644 2664 4 (POTENTIAL).  
FT DOMAIN 2665 2681 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 2682 2702 5 (POTENTIAL).  
FT DOMAIN 2703 2725 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 2726 2746 6 (POTENTIAL).  
FT DOMAIN 2747 2753 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 2754 2774 7 (POTENTIAL).  
FT DOMAIN 2775 3312 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 336 433 CADHERIN 1.  
FT DOMAIN 434 545 CADHERIN 2.  
FT TRANSMEM 546 651 CADHERIN 3.  
FT DOMAIN 652 756 CADHERIN 4.  
FT TRANSMEM 757 858 CADHERIN 5.  
FT DOMAIN 859 961 CADHERIN 6.  
FT TRANSMEM 962 1067 CADHERIN 7.  
FT DOMAIN 1068 1169 CADHERIN 8.  
FT TRANSMEM 1170 1265 CADHERIN 9.

FT DOMAIN 1375 1433 EGF-LIKE 1, CALCIUM-BINDING.  
FT DOMAIN 1435 1471 EGF-LIKE 2, CALCIUM-BINDING.  
FT DOMAIN 1475 1514 EGF-LIKE 3, CALCIUM-BINDING.  
FT DOMAIN 1515 1719 LAMININ G-LIKE 1.  
FT DOMAIN 1722 1758 EGF-LIKE 4, CALCIUM-BINDING.  
FT DOMAIN 1764 1944 LAMININ G-LIKE 2.  
FT DOMAIN 1946 1982 EGF-LIKE 5, CALCIUM-BINDING.  
FT DOMAIN 2021 2053 EGF-LIKE 6, CALCIUM-BINDING.  
FT DOMAIN 2055 2090 EGF-LIKE 7, CALCIUM-BINDING.  
FT DOMAIN 2096 2131 EGF-LIKE 8, CALCIUM-BINDING.  
FT DOMAIN 2477 2529 LAMININ EGF-LIKE.  
FT DISULFID 1379 1390 GPS.  
FT DISULFID 1384 1421 BY SIMILARITY.  
FT DISULFID 1432 1432 BY SIMILARITY.  
FT DISULFID 1439 1450 BY SIMILARITY.  
FT DISULFID 1444 1459 BY SIMILARITY.  
FT DISULFID 1461 1470 BY SIMILARITY.  
FT DISULFID 1479 1490 BY SIMILARITY.  
FT DISULFID 1484 1500 BY SIMILARITY.  
FT DISULFID 1502 1513 BY SIMILARITY.  
FT DISULFID 1726 1737 BY SIMILARITY.  
FT DISULFID 1731 1746 BY SIMILARITY.  
FT DISULFID 1748 1757 BY SIMILARITY.  
FT DISULFID 1950 1961 BY SIMILARITY.  
FT DISULFID 1955 1970 BY SIMILARITY.  
FT DISULFID 1972 1981 BY SIMILARITY.  
FT DISULFID 1985 1996 BY SIMILARITY.  
FT DISULFID 1990 2008 BY SIMILARITY.  
FT DISULFID 2010 2019 BY SIMILARITY.  
FT DISULFID 2027 2040 BY SIMILARITY.  
FT DISULFID 2042 2052 BY SIMILARITY.  
FT DISULFID 2059 2074 BY SIMILARITY.  
FT DISULFID 2061 2077 BY SIMILARITY.  
FT DISULFID 2079 2089 BY SIMILARITY.  
FT MOD\_RES 1963 1963 HYDROXYLATION (POTENTIAL).  
FT CARBOHYD 632 632 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 847 847 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1182 1182 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1222 1222 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1317 1317 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1327 1327 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1649 1649 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1713 1713 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1770 1770 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 2053 2053 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 2177 2177 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 2196 2196 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 2386 2386 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 2474 2474 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 2506 2506 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CONFLICT 2158 2158 G -> GLRGAG (IN REF. 2).  
SQ SEQUENCE 3312 AA; 358251 MW; BEC208703651AA5 CRC64;

Query Match 14.1%; Score 81.5; DB 1; Length 3312;  
Best Local Similarity 34.4%; Pred. No. 20;  
Matches 31; Conservative 6; Mismatches 38; Indels 15; Gaps 5;

QY 16 AMLAAQRRVPRAAEVPGAQGGQGGPRGEEAPRGVRMA-----VPLLRRMEGAPAGPGR 70  
Db 220 ATTSGAERTAPRRNCLPGASG-SGPE-LDSAPTARTAPASGSAPRESRTAPEAPKMR 277  
QY 71 TAACFSCYTRCLSRPRWRKRSWSAGSCPGMP 100  
Db 278 SRGLFRCL--RFLPQRPGRPR-----PGLP 299

RESULT 4  
SRRL\_HUMAN  
ID SRRL\_HUMAN STANDARD; PRT; 339 AA.  
AC Q9UH36;  
DT 15-SEP-2003 (Rel. 42, Created)  
DT 15-SEP-2003 (Rel. 42, Last sequence update)

15-SEP-2003 (Rel. 42, Last annotation update)  
 SRRI-like protein.  
 DE Homo sapiens (Human).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20057165; PubMed=10591208;  
 RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,  
 RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,  
 RA Baguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,  
 RA Bird C.P., Blakey S.E., Bridgman A.M., Buck D., Burgess J.,  
 RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,  
 RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,  
 RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,  
 RA Dhani P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,  
 RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,  
 RA Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,  
 RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,  
 RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,  
 RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,  
 RA Martyn I.D., Mashreghi-Mohammadi M., Milne S.A., Mortimore B.J.,  
 RA Mcclay J., McLaren S., McMurray A.A., Pearson D., Phillimore B.J.C.T.,  
 RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,  
 RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,  
 RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,  
 RA Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,  
 RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,  
 RA Williams L., Williams S., Williamson H., Wilmer T.E., Wilming L.,  
 RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,  
 RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,  
 RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,  
 RA Roe B.A., Chen F., Chu Y., Crabtree J., Deschamps S., Do A., Do T.,  
 RA Dorman A., Fang F., Fu Y., Hua A., Kenton S., Lai H., Lao H.I.,  
 RA Lewis J., Lewis S., Lin S.-P., Loh P., Mala J.E., Nguyen T., Pan H.,  
 RA Phan S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,  
 RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,  
 RA Zhan M., Zhang G., Chisoe S., Murray J., Miller N., Mix P.,  
 RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,  
 RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,  
 RA Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Nohlfing T.,  
 RA Scheet P., Walker C., Wamsley A., Wohldmann P., Pepin K., Neilson J.,  
 RA Korf I., Bedell J.A., Hillier L., Mardis E., Waterston R., Wilson R.,  
 RA Emanuel B.S., Shaikh T., Kurahashi H., Saitta S., Budarf M.L.,  
 RA McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E., Edelman L.,  
 RA Kim U.J., Shizuya H., Simon M.I., Dumanski J.P., Peyrard M., Kedra D.,  
 RA Seroussi E., Fransson I., Tapia I., Bruder C.E., O'Brien K.P.,  
 RA Wilkison P., Bodenteich A., Hartman K., Hu X., Khan A.S., Lane L.,  
 RA Tilahun Y., Wright H.;  
 RL "the DNA sequence of human chromosome 22.";  
 RL Nature 402:489-495(1999)  
 CC -!- FUNCTION: May be involved in a circadian clock input pathway (By  
 CC similarity).  
 CC -!- SIMILARITY: Belongs to the SRRI family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: Z99714; CAB62929.1;  
 DR Hypothetical protein: Repeat.  
 KW DOMAIN 30 50 3 X TANDEM REPEAT OF R-E-A-A-P-R-G.  
 FT REPEAT 30 36 1.  
 FT REPEAT 37 43 2.  
 FT REPEAT 44 50 3.  
 FT DOMAIN 2 8 POLY-ALA.  
 FT DOMAIN 162 166 POLY-LEU.  
 FT SEQUENCE 339 AA; 38573 MW; 9DDF4C81ED940468 CRC64;  
 SQ

Query Match 14.0%; Score 81; DB 1; Length 339;  
 Best Local Similarity 32.4%; Pred. NO. 3; 3; Gaps 5;  
 Matches 33; Conservative 5; Mismatches 32; Indels 32; Gaps 5;  
 QY 3 MAQEAALAFMAQAGMLAAQER---RYPRAAEVFGAQQGQGRGREAAPRGVMAVPLLR 59  
 Db 1 MAAAAAALLESWQAAPRRKRRAA-----APRGREAAPRG----- 43  
 QY 60 MEGAPAGPGGRTAACFSCSTSRCLSRPWRKRS-----WSA 93  
 Db 44 REAAPRGP-----EAEFESDSGVVLRRIWEAKDLFISDFWSS 81  
 RESULT 5  
 ID ICP4\_HSV11 STANDARD; PRT; 1298 AA.  
 AC P08392;  
 DT 01-AUG-1988 (Rel. 08, Created)  
 DT 01-AUG-1988 (Rel. 08, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Trans-acting transcriptional protein ICP4 (Transcriptional activator  
 DE IE175) (Alpha-4 protein).  
 GN ICP4 OR IE175 OR RSL.  
 OS Herpes simplex virus (type 1 / strain 17).  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Alphaherpesvirinae; Simplexvirus.  
 OX NCBI\_TaxID=10299;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88274327; PubMed=2839594;  
 RA McGeoch D.J., Dairymple M.A., Davison A.J., Dolan A., Frame M.C.,  
 RA McNab D., Perry L.J., Scott J.E., Taylor P.,  
 RT "The complete DNA sequence of the long unique region in the genome of  
 RT herpes simplex virus type 1.";  
 RL J. Gen. Virol. 69:1531-1574(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86148504; PubMed=3005980;  
 RA McGeoch D.J., Dolan A., Donald S., Brauer D.H.K.;  
 RT "Complete DNA sequence of the short repeat region in the genome of  
 RT herpes simplex virus type 1.";  
 RL Nucleic Acids Res. 14:1727-1745(1986).  
 RN [3]  
 RP DNA-BINDING DOMAIN.  
 RX MEDLINE=90174974; PubMed=2155403;  
 RA Wu C.-L., Wilcox K.W.;  
 RT "Codons 262 to 490 from the herpes simplex virus ICP4 gene are  
 RT sufficient to encode a sequence-specific DNA binding protein.";  
 RL Nucleic Acids Res. 18:531-538(1990).  
 RN [4]  
 RP INFLUENCE OF PHOSPHORYLATION ON FUNCTION.  
 RX MEDLINE=91122047; PubMed=1846804;  
 RA Papavasiliou A.G., Wilcox K.W., Silverstein S.J.;  
 RT "The interaction of ICP4 with cell/infected-cell factors and its  
 RT state of phosphorylation modulate differential recognition of leader  
 RT sequences in herpes simplex virus DNA.";  
 RL EMBO J. 10:397-406(1991).  
 CC -!- FUNCTION: THIS IE PROTEIN IS A MULTIFUNCTIONAL PROTEIN CAPABLE  
 CC OF MIGRATING TO THE NUCLEUS, BINDING TO DNA, TRANS-ACTIVATING  
 CC OTHER VIRAL GENES, AND AUTOREGULATING ITS OWN SYNTHESIS. BINDING  
 CC OF ICP4 TO ALPHA GENES CAUSES REPRESSION OF THEIR EXPRESSION.  
 CC ICP4 ALSO ACTS AS A TRANSCRIPTIONAL ACTIVATOR OF BETA AND GAMMA  
 CC GENES. ICP4 BINDS WITH HIGH AFFINITY TO THE SEQUENCE 5'-ATCGTC-3'.  
 CC -!- SUBUNIT: Homodimer.  
 CC -!- SUBCELLULAR LOCATION: NUCLEUS OF INFECTED CELLS.  
 CC -!- PTM: ICP4 IS ADP-RIBOSYLATED.  
 CC -!- PTM: THE LONG STRETCH OF SER IS A MAJOR SITE OF PHOSPHORYLATION.  
 CC ONLY THE PHOSPHORYLATED FORMS OF ICP4 ARE CAPABLE OF INTERACTING  
 CC WITH BETA OR GAMMA GENES.  
 CC -!- SIMILARITY: BELONGS TO THE ICP4/IE140/IE180 FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration





```
QY 88 KRSWSAGSCPMPLSPDQGR 108
Db 148 VPVWTAASCALLAHIRTDLSR 168

RESULT 8
ID UFO3_MAIZE STANDARD; PRT; 471 AA.
AC P16166; Q9ATX0;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Flavonol 3-O-glucosyltransferase (EC 2.4.1.91) (UDP-glucose flavonoid
DE 3-O-glucosyltransferase) (Bronze-1) (Bz-McC allele).
DE BZ1 OR UGT71A1.
GN Zea mays (Maize).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
SEQUENCE FROM N.A.
RP Furtak D., Schiefelbein J.W., Johnston F., Nelson O.E. Jr.;
RA "Sequence comparisons of 3 wild-type bronze-1 alleles from Zea mays.";
RL Plant Mol. Biol. 11:473-481(1988).
RN [2]
SEQUENCE FROM N.A.
RP MEDLINE=88284304; PubMed=3396861;
RA Ralston E.J., English J.J., Dooner H.K.;
RT "Sequence of three bronze alleles of maize and correlation with the
RT genetic fine structure.";
RL Genetics 119:185-197(1988).
CC -!- FUNCTION: IN THE PRESENCE OF OTHER NECESSARY COLOR FACTORS, THIS
CC GLYCOSYLATION REACTION ALLOWS THE ACCUMULATION OF ANTHOCYANIN
CC PIGMENTS.
CC -!- CATALYTIC ACTIVITY: UDP-glucose + a flavonol = UDP + flavonol
CC 3-O-D-glucoside.
CC -!- PATHWAY: Anthocyanin biosynthesis.
CC -!- SIMILARITY: BELONGS TO THE UDP-GLYCOSYLTRANSFERASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X13500; CAA31855.1;
CC EMBL; AF391808; AAK73112.1;
CC PIR; S01052; S01052.
CC MaizeDB; 13885;
CC InterPro; IPR002213; UDP_gluco_trans.
CC Pfam; PF00201; UDPGT; 1.
CC PROSITE; PS00375; UDPGT; 1.
CC Transferrase; Glycosyltransferase.
CC SEQUENCE 471 AA; 48769 MW; 6234FDS9219AF534 CRC64;

Query Match 13.6%; Score 78.5; DB 1; Length 471;
Best Local Similarity 24.6%; Pred. No. 7.1;
Matches 34; Conservative 14; Mismatches 57; Indels 33; Gaps 5;

QY 1 MLMAQEAALAFUMA-OGAML-----AAQERRVPRAAEVPGAGGOGPRGEEAPRGVMA- 53
Db 28 LLSTARALAAAAAASGATLSFLSTASSLAQLRKASSASAGHGLPGLNLFVEVDPGAPAAE 87
QY 54 --VPLLRM-----EGAPAGPGGRTAACF-----SCTSRCLSRP 87
Db 88 ETVVPVPMQQLFMEAAEAGGVKAWLEAARAAGGARTVCVVDFAFWPAAADAASAGAPW 147
QY 88 KRSWSAGSCPMPLSPD 105
Db 148 VPVWTAASCALLAHIRTD 165

RESULT 10
QY 88 KRSWSAGSCPMPLSPDQGR 108
Db 148 VPVWTAASCALLAHIRTDLSR 168

RESULT 9
ID UFO3_MAIZE STANDARD; PRT; 471 AA.
AC P16167;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Flavonol 3-O-glucosyltransferase (EC 2.4.1.91) (UDP-glucose flavonoid
DE 3-O-glucosyltransferase) (Bronze-1) (Bz-W22 allele).
DE BZ1 OR UGT71A1.
GN Zea mays (Maize).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
SEQUENCE FROM N.A.
RP Furtak D., Schiefelbein J.W., Johnston F., Nelson O.E. Jr.;
RA "Sequence comparisons of 3 wild-type bronze-1 alleles from Zea mays.";
RL Plant Mol. Biol. 11:473-481(1988).
RN [2]
SEQUENCE FROM N.A.
RP MEDLINE=88284304; PubMed=3396861;
RA Ralston E.J., English J.J., Dooner H.K.;
RT "Sequence of three bronze alleles of maize and correlation with the
RT genetic fine structure.";
RL Genetics 119:185-197(1988).
CC -!- FUNCTION: IN THE PRESENCE OF OTHER NECESSARY COLOR FACTORS, THIS
CC GLYCOSYLATION REACTION ALLOWS THE ACCUMULATION OF ANTHOCYANIN
CC PIGMENTS.
CC -!- CATALYTIC ACTIVITY: UDP-glucose + a flavonol = UDP + flavonol
CC 3-O-D-glucoside.
CC -!- PATHWAY: Anthocyanin biosynthesis.
CC -!- SIMILARITY: BELONGS TO THE UDP-GLYCOSYLTRANSFERASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X13502; CAA31857.1;
CC EMBL; X07937; CAA30760.1;
CC PIR; S01037; S01037.
CC MaizeDB; 13885;
CC InterPro; IPR002213; UDP_gluco_trans.
CC Pfam; PF00201; UDPGT; 1.
CC PROSITE; PS00375; UDPGT; 1.
CC Transferrase; Glycosyltransferase.
CC SEQUENCE 471 AA; 48673 MW; F417020B78366C01 CRC64;

Query Match 13.6%; Score 78.5; DB 1; Length 471;
Best Local Similarity 24.6%; Pred. No. 7.1;
Matches 34; Conservative 14; Mismatches 57; Indels 33; Gaps 5;

QY 1 MLMAQEAALAFUMA-OGAML-----AAQERRVPRAAEVPGAGGOGPRGEEAPRGVMA- 53
Db 28 LLSTARALAAAAAASGATLSFLSTASSLAQLRKASSASAGHGLPGLNLFVEVDPGAPAAE 87
QY 54 --VPLLRM-----EGAPAGPGGRTAACF-----SCTSRCLSRP 87
Db 88 ETVVPVPMQQLFMEAAEAGGVKAWLEAARAAGGARTVCVVDFAFWPAAADAASAGAPW 147
QY 88 KRSWSAGSCPMPLSPD 105
Db 148 VPVWTAASCALLAHIRTD 165

RESULT 10
```



RESULT 12  
 ID CAL3\_HUMAN STANDARD; PRT; 1466 AA.  
 AC P02461; Q15112;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Collagen alpha 1(III) chain precursor.  
 GN COL3A1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Skin fibroblast;  
 RX MEDLINE=89350838; PubMed=2764886;  
 RA Ala-Kokko L., Kontusaari S., Baldwin C.T., Kuivaniemi H.,  
 RA Prockop D.J.;  
 RT "Structure of cDNA clones coding for the entire prepro alpha 1 (III)  
 RT chain of human type III procollagen. Differences in protein structure  
 RT from type I procollagen and conservation of codon preferences.";  
 RL Biochem. J. 260:509-516(1989).  
 RN [2]  
 RP SEQUENCE OF 149-1225 FROM N.A.  
 RX MEDLINE=89386015; PubMed=2780304;  
 RA Janczko R.A., Ramirez F.;  
 RT "Nucleotide and amino acid sequences of the entire human alpha 1  
 RT (III) collagen.";  
 RL Nucleic Acids Res. 17:6742-6742(1989).  
 RN [3]  
 RP SEQUENCE OF 168-398.  
 RX MEDLINE=77134724; PubMed=557335;  
 RA Seyer J.M., Kang A.H.;  
 RT "Covalent structure of collagen: amino acid sequence of cyanogen  
 RT bromide peptides from the amino-terminal segment of type III collagen  
 RT of human liver.";  
 RL Biochemistry 16:1158-1164(1977).  
 RN [4]  
 RP REVISIONS.  
 RA Seyer J.M.;  
 RL Submitted (DEC-1977) to the PIR data bank.  
 RN [5]  
 RP SEQUENCE OF 399-727.  
 RX MEDLINE=79000343; PubMed=687591;  
 RA Seyer J.M., Kang A.H.;  
 RT "Covalent structure of collagen: amino acid sequence of five  
 RT consecutive CNBr peptides from type III collagen of human liver.";  
 RL Biochemistry 17:3404-3411(1978).  
 RN [6]  
 RP SEQUENCE OF 728-964.  
 RX MEDLINE=80198282; PubMed=6246925;  
 RA Seyer J.M., Mainardi C., Kang A.H.;  
 RT "Covalent structure of collagen: amino acid sequence of alpha 1  
 RT (III)-CB5 from type III collagen of human liver.";  
 RL Biochemistry 19:1583-1589(1980).  
 RN [7]  
 RP SEQUENCE OF 950-1466 FROM N.A.  
 RX MEDLINE=88189827; PubMed=3357782;  
 RA Mankoo B.S., Daigleish R.;  
 RT "Human pro alpha 1(III) collagen: cDNA sequence for the 3' end.";  
 RL Nucleic Acids Res. 16:2337-2337(1988).  
 RN [8]  
 RP REVISION TO 1184.  
 RX MEDLINE=89083436; PubMed=3211760;  
 RA Molyneux K., Daigleish R.;  
 RT "Human type III collagen 'variant' is a cDNA cloning artefact.";  
 RL Nucleic Acids Res. 16:11833-11833(1988).  
 RN [9]  
 RP SEQUENCE OF 1065-1466 FROM N.A.  
 RX MEDLINE=85087944; PubMed=6096827;  
 RA Loidi H.R., Brinker J.M., May M., Pihlajaniemi T., Morrow S.,  
 RA Rosenbloom J., Myers J.C.;  
 RT "Molecular cloning and carboxyl-propeptide analysis of human type III  
 RT procollagen.";  
 RL Nucleic Acids Res. 12:9383-9394(1984).  
 RN [10]  
 RP SEQUENCE OF 965-1200.  
 RX MEDLINE=81208139; PubMed=7016180;  
 RA Seyer J.M., Kang A.H.;  
 RT "Covalent structure of collagen: amino acid sequence of alpha  
 RT 1(III)-CB9 from type III collagen of human liver.";  
 RL Biochemistry 20:2621-2627(1981).  
 RN [11]  
 RP SEQUENCE OF 1176-1466 FROM N.A.  
 RX MEDLINE=85157600; PubMed=2579949;  
 RA Chu M.-L., Weil D., de Wet W.J., Bernard M.P., Sippola M., Ramirez F.;  
 RT "Isolation of cDNA and genomic clones encoding human pro-alpha 1  
 RT (III) collagen. Partial characterization of the 3' end region of the  
 RT gene.";  
 RL J. Biol. Chem. 260:4357-4363(1985).  
 RN [12]  
 RP SEQUENCE OF 1161-1200 FROM N.A.  
 RX MEDLINE=86187804; PubMed=3754462;  
 RA Miskulin M., Daigleish R., Kluge-Beckerman B., Rennard S.I.,  
 RA Tolstoshev P., Brantly M., Crystal R.G.;  
 RT "Human type III collagen gene expression is coordinately modulated  
 RT with the type I collagen genes during fibroblast growth.";  
 RL Biochemistry 25:1408-1413(1986).  
 RN [13]  
 RP SEQUENCE OF 1-170 FROM N.A.  
 RC TISSUE-Placenta;  
 RX MEDLINE=88303360; PubMed=3405773;  
 RA Toman D., Ricca G., de Crombrughe B.;  
 RT "Nucleotide sequence of a cDNA coding for the amino-terminal region  
 RT of human prepro alpha 1(III) collagen.";  
 RL Nucleic Acids Res. 16:7201-7201(1988).  
 RN [14]  
 RP SEQUENCE OF 1-176 FROM N.A.  
 RX MEDLINE=89378752; PubMed=2777083;  
 RA Benson-Chanda V., Su M.W., Weil D., Chu M.-L., Ramirez F.;  
 RT "Cloning and analysis of the 5' portion of the human type-III  
 RT procollagen gene (COL3A1).";  
 RL Gene 78:255-265(1989).  
 RN [15]  
 RP REVIEW ON VARIANTS.  
 RX MEDLINE=97255959; PubMed=9101290;  
 RA Kuivaniemi H., Tromp G., Prockop D.J.;  
 RT "Mutations in fibrillar collagens (types I, II, III, and XI), fibril-  
 RT associated collagen (type IX), and network-forming collagen (type X)  
 RT cause a spectrum of diseases of bone, cartilage, and blood vessels.";  
 RL Hum. Mutat. 9:300-315(1997).  
 RN [16]  
 RP VARIANT AORTIC ANEURYSM ARG-303, AND VARIANT THR-668.  
 RX MEDLINE=93293988; PubMed=8514866;  
 RA Tromp G., Wu Y., Prockop D.J., Madhathari S.L., Kleinert C.,  
 RA Earley J.J., Zhuang J., Noerregaard O., Darling R.C., Abbott W.M.,  
 RA Cole C.W., Jaakkola P., Ryyanen M., Pearce W.H., Yao J.S.-T.,  
 RA Majamaa K., Smellens S.N., Gatalica Z., Ferrell R.E., Jimenez S.A.,  
 RA Jackson C.E., Michels V.V., Kaye M., Kuivaniemi H.;  
 RT "Sequencing of cDNA from 50 unrelated patients reveals that mutations  
 RT in the triple-helical domain of type III procollagen are an  
 RT infrequent cause of aortic aneurysms.";  
 RL J. Clin. Invest. 91:2539-2545(1993).  
 RN [17]  
 RP VARIANT THR-698.  
 RX MEDLINE=91045136; PubMed=2235526;  
 RA Zafarullah K., Kleinert C., Tromp G., Kuivaniemi H., Kontusaari S.,  
 RA Wu Y., Ganguly A., Prockop D.J.;  
 RT "G to A polymorphism in exon 3 of the COL3A1 gene.";  
 RL Nucleic Acids Res. 18:6180-6180(1990).  
 RN [18]  
 RP VARIANT AORTIC ANEURYSM ARG-786.  
 RX MEDLINE=91056145; PubMed=2243125;  
 RA Kontusaari S., Tromp G., Kuivaniemi H., Romanic A.M., Prockop D.J.;

RT "A mutation in the gene for type III procollagen (COL3A1) in a family  
 RL with aortic aneurysms.";  
 RN J. Clin. Invest. 86:1465-1473(1990).  
 RP [19]  
 RP VARIANT EDS-IV ARG-828.  
 RX MEDLINE-94016385; PubMed-8411057;  
 RA Richards A.J., Narcisi P., Lloyd J.C., Ferguson C., Pope F.M.;  
 RT "The substitution of glycine 661 by arginine in type III collagen  
 RT produces mutant molecules with different thermal stabilities and  
 RT causes Ehlers-Danlos syndrome type IV.";  
 RL J. Med. Genet. 30:690-693(1993).  
 RN [20]  
 RP VARIANT EDS-IV SER-957.  
 RX MEDLINE-89109135; PubMed-2492273;  
 RA Tromp G., Kuivaniemi H., Shikata H., Prockop D.J.;  
 RT "A single base mutation that substitutes serine for glycine 790 of  
 RT the alpha 1 (III) chain of type III procollagen exposes an arginine  
 RT and causes Ehlers-Danlos syndrome IV.";  
 RL J. Biol. Chem. 264:1349-1352(1989).  
 RN [21]  
 RP VARIANT EDS-IV VAL-960.  
 RX MEDLINE-95268429; PubMed-7749417;  
 RA Tromp G., de Paeppe A., Nuytinck L., Madhathari S.L., Kuivaniemi H.;  
 RT "Substitution of valine for glycine 793 in type III procollagen in  
 RT Ehlers-Danlos syndrome type IV.";  
 RL Hum. Mutat. 5:179-181(1995).  
 RN [22]  
 RP VARIANT EDS-IV GLU-1014.  
 RX MEDLINE-92316511; PubMed-1352273;  
 RA Richards A.J., Ward P.N., Narcisi P., Nicholls A.C., Lloyd J.C.,  
 RA Pope F.M.;  
 RT "A single base mutation in the gene for type III collagen (COL3A1)  
 RT converts glycine 847 to glutamic acid in a family with Ehlers-Danlos  
 RT syndrome type IV. An unaffected family member is mosaic for the  
 RT mutation.";  
 RL Hum. Genet. 89:414-418(1992).  
 RN [23]  
 RP VARIANT EDS-IV ASP-1050.  
 RX MEDLINE-90037070; PubMed-2808425;  
 RA Tromp G., Kuivaniemi H., Stolle C.A., Pope F.M., Prockop D.J.;  
 RT "Single base mutation in the type III procollagen gene that converts  
 RT the codon for glycine 883 to aspartate in a mild variant of  
 RT Ehlers-Danlos syndrome IV.";  
 RL J. Biol. Chem. 264:19313-19317(1989).  
 RN [24]  
 RP VARIANT EDS-IV VAL-1077.  
 RX MEDLINE-91374480; PubMed-1895316;  
 RA Richards A.J., Lloyd J.C., Ward P.N., de Paeppe A., Narcisi P.,  
 RA Pope F.M.;  
 RT "Characterisation of a glycine to valine substitution at amino acid  
 RT position 910 of the triple helical region of type III collagen in a  
 RT patient with Ehlers-Danlos syndrome type IV.";  
 RL J. Med. Genet. 28:458-463(1991).  
 RN [25]  
 RP VARIANT EDS-IV GLU-1173.  
 RX MEDLINE-93022543; PubMed-1357232;  
 RA Johnson P.H., Richards A.J., Pope F.M., Hopkinson D.A.;

Query Match 13.5%; Score 78; DB 1; Length 1466;  
 Best Local Similarity 30.4%; Pred. No. 20;  
 Matches 34; Conservative 5; Mismatches 33; Indels 40; Gaps 6;  
 QY 13 AQAAMLAQAQRRVPP--RAAEVP-GAQQQGGPRGEEA--PRGVRMAV-----PL 56  
 Db 475 ANGLPCAAGERGAPGFRGAGNGIPGKGPAGERGAPGAPGAGGPRGDPGPG 534  
 QY 57 LRMGAPAGPGGRTAACFSCTSRCLSRPWRKWSAGSCPGMPLSPDQGR 108  
 Db 535 MRGMPSGPGPG-----SDGR-PGPPGSGGESGR 562

RESULT 13  
 CTG1\_HUMAN

ID CTG1\_HUMAN STANDARD; PRT; 180 AA.  
 AC P78358;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cancer/testis antigen 1 (Autoimmunogenic cancer/testis antigen NY-ESO-1).  
 GN CTAG1 OR CTAG.  
 GE Homo sapiens (Human).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 OX [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE-97203161; PubMed-9050879;  
 RA Chen Y.-T., Scanlan M.J., Sahin U., Tuerci O., Gure A.O., Tsang S.,  
 RA Williamson B., Stockert E., Pfeundscher M., Old L.J.;  
 RT "A testicular antigen aberrantly expressed in human cancers detected  
 RT by autologous antibody screening.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:1914-1918(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Melanoma;  
 RX MEDLINE-98289662; PubMed-9626360;  
 RA Lethe B., Lucas S., Michaux L., de Smet C., Godelaine D., Serrano A.,  
 RA de Plaen E., Boon T.;  
 RT "LAGE-1, a new gene with tumor specificity.";  
 RL Int. J. Cancer 76:903-908(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-98430682; PubMed-9759882;  
 RA Wang R.-F., Johnston S.L., Zeng G., Topalian S.L.,  
 RA Schwartzentruber D.J., Rosenberg S.A.;  
 RT "A breast and melanoma-shared tumor antigen: T cell responses to  
 RT antigenic peptides translated from different open reading frames.";  
 RL J. Immunol. 161:3596-3606(1998).  
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN TESTIS AND OVARY AND IN A WIDE  
 CC VARIETY OF CANCERS. DETECTED IN UTERINE MYOMETRIUM.  
 CC -!- SIMILARITY: BELONGS TO THE CTAG FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U87459; AAB49693.1; -;  
 DR EMBL; AJ003149; CA005908.1; -;  
 DR EMBL; AF038567; AAD05202.1; -;  
 DR Genbank; HGNC:2491; CTAG1.  
 DR MIM; 300156; -;  
 KW Transmembrane; Antigen.  
 FT DOMAIN 5 82 GLY-RICH.  
 FT TRANSMEM 156 172 POTENTIAL.  
 SQ SEQUENCE 180 AA; 17992 MW; B122C5C2C8BE1569 CRC64;  
 Query Match 13.3%; Score 77; DB 1; Length 180;  
 Best Local Similarity 36.1%; Pred. No. 4.3;  
 Matches 22; Conservative 4; Mismatches 19; Indels 16; Gaps 2;  
 QY 33 GAQQQGGPRG---REEAPRGVRMAVPLLRMRMEGAPAGPGGRTAACFSCTSRCLSRPWK 88  
 Db 38 GATGGRGPRGAGAAASGPG-----GAPRGHGAASGLNGCCRCGARGPES 85  
 QY 89 R 89  
 Db 86 R 86  
 RESULT 14  
 CA26\_HUMAN

ID	CA26_HUMAN	STANDARD;	PRT; 1019 AA.	RT	"Type VI collagen mutations in Bethlem myopathy, an autosomal dominant
AC	P12110; Q13909; Q13910; Q13911; Q14049; Q16259; Q16597; Q9UML3;			RT	myopathy with contractures.";
DT	01-OCT-1989 (Rel. 12, Created)			RL	Nat. Genet. 14:113-115(1996).
DT	28-FEB-2003 (Rel. 41, Last sequence update)			RN	[9]
DT	28-FEB-2003 (Rel. 41, Last annotation update)			RX	VARIANT BM ASN-621.
DE	Collagen alpha 2(VI) chain precursor.			RX	MEDLINE-21853823; PubMed-11865138;
GN	COL6A2.			RA	Scacheri P.C., Gillanders E.M., Subramony S.H., Vedanarayanan V.,
OS	Homo sapiens (Human).			RA	Scacheri P.C., Gillanders E.M., Subramony S.H., Vedanarayanan V.,
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			RA	Scacheri P.C., Gillanders E.M., Subramony S.H., Vedanarayanan V.,
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			RA	Scacheri P.C., Gillanders E.M., Subramony S.H., Vedanarayanan V.,
OX	NCBI_TaxID=9606;			RA	Scacheri P.C., Gillanders E.M., Subramony S.H., Vedanarayanan V.,
RP	SEQUENCE FROM N.A., AND REVISIONS.			RT	"Novel mutations in collagen VI genes: expansion of the Bethlem
RC	TISSUE=Fibroblast, and Placenta;			RT	myopathy phenotype.";
RA	Chu M.-L.;			RL	Neurology 58:593-602(2002).
RL	Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.			CC	!- FUNCTION: COLLAGEN VI ACTS AS A CELL-BINDING PROTEIN. ALPHA 1(VI),
RN	SEQUENCE OF 1-255 AND 591-1019 FROM N.A.			CC	!- SUBUNIT: TRIMERS COMPOSED OF THREE DIFFERENT CHAINS: ALPHA 1(VI),
RP	TISSUE=Fibroblast;			CC	ALPHA 2(VI), AND ALPHA 3(VI).
RX	MEDLINE-90005396; PubMed-2551668;			CC	!- ALTERNATIVE PRODUCTS:
RA	Chu M.-L., Pan T.-C., Conway D., Kuo H.J., Glanville R.W., Timpl R.,			CC	Event-Alternative splicing; Named isoforms=3;
RA	Mann K., Deutzmann R.;			CC	Name=2C2;
RT	"Sequence analysis of alpha 1(VI) and alpha 2(VI) chains of human type			CC	IsoId=P12110-1; Sequence=VSP_001163, VSP_001164;
RT	VI collagen reveals internal triplication of globular domains similar			CC	IsoId=P12110-2; Sequence=VSP_001163, VSP_001164;
RT	to the A domains of von Willebrand factor and two alpha 2(VI) chain			CC	Name=2C2A;
RT	variants that differ in the carboxy terminus.";			CC	IsoId=P12110-3; Sequence=VSP_001161, VSP_001162;
RL	EMBO J. 8:1939-1946(1989).			CC	!- PTM: Prolines at the third position of the tripeptide repeating
RN	[3]			CC	unit (G-X-Y) are hydroxylated in some or all of the chains.
RP	SEQUENCE OF 238-299 FROM N.A.			CC	!- DISEASE: Defects in COL6A2 are a cause of Bethlem myopathy (BM)
RX	MEDLINE-88029444; PubMed-3665927;			CC	[MIM:158810]. BM is a rare autosomal dominant proximal myopathy
RA	Chu M.-L., Mann K., Deutzmann R., Pribula-Conway D.,			CC	characterized by early childhood onset (complete penetrance by the
RA	Hsu-Chen C.-C., Bernard M.P., Timpl R.;			CC	age of 5) and joint contractures most frequently affecting the
RT	"Characterization of three constituent chains of collagen type VI by			CC	elbows and ankles.
RT	peptide sequences and cDNA clones.";			CC	!- DISEASE: Defects in COL6A2 are a cause of Ullrich congenital
RL	Eur. J. Biochem. 168:309-317(1987).			CC	muscular dystrophy (UCMD) [MIM:254090]; also known as Ullrich
RN	[4]			CC	sclerotic muscular dystrophy, an autosomal recessive congenital
RP	SEQUENCE OF 246-590 FROM N.A.			CC	myopathy. UCMD is characterized by muscle weakness and multiple
RX	MEDLINE-92112205; PubMed-1765372;			CC	joint contractures, generally noted at birth or early infancy. The
RA	Saitta B., Wang Y.M., Renkart L., Zhang R.Z., Pan T.-C., Timpl R.,			CC	clinical course is more severe than in Bethlem myopathy.
RA	Chu M.-L.;			CC	!- SIMILARITY: Contains 3 WFMA domains.
RT	"The exon organization of the triple-helical coding regions of the			CC	-----
RT	human alpha 1(VI) and alpha 2(VI) collagen genes is highly similar.";			CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
RL	Genomics 11:145-153(1991).			CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
RN	[5]			CC	the European Bioinformatics Institute. There are no restrictions on its
RP	SEQUENCE OF 256-590 FROM N.A.			CC	use by non-profit institutions as long as its content is in no way
RX	MEDLINE-89066644; PubMed-3158591;			CC	modified and this statement is not removed. Usage by and for commercial
RA	Chu M.-L., Conway D., Pan T.-C., Baldwin C., Mann K., Deutzmann R.,			CC	entities requires a license agreement (See <a href="http://www.lsb-sib.ch/announce/">http://www.lsb-sib.ch/announce/</a>
RA	Timpl R.;			CC	or send an email to <a href="mailto:license@lsb-sib.ch">license@lsb-sib.ch</a> ).
RT	"Amino acid sequence of the triple-helical domain of human collagen			CC	-----
RT	type VI.";			CC	EMBL; AY029208; AAA52056.2; -
RL	J. Biol. Chem. 263:18601-18606(1988).			CC	EMBL; X06195; CAA29556.1; -
RN	[6]			CC	EMBL; S75462; AAB20836.1; -
RP	SEQUENCE OF 316-359 FROM N.A.			CC	EMBL; S75425; AAB20836.1; JOINED.
RC	TISSUE=Placenta;			CC	EMBL; S75428; AAB20836.1; JOINED.
RX	MEDLINE-88161046; PubMed-3348212;			CC	EMBL; S75430; AAB20836.1; JOINED.
RA	Well D., Mattei M.-G., Passage E., N'Guyen V.C., Pribula-Conway D.,			CC	EMBL; S75432; AAB20836.1; JOINED.
RA	Mann K., Deutzmann R., Timpl R., Chu M.-L.;			CC	EMBL; S75434; AAB20836.1; JOINED.
RT	"Cloning and chromosomal localization of human genes encoding the			CC	EMBL; S75436; AAB20836.1; JOINED.
RT	three chains of type VI collagen.";			CC	EMBL; S75438; AAB20836.1; JOINED.
RL	Am. J. Hum. Genet. 42:435-445(1988).			CC	EMBL; S75440; AAB20836.1; JOINED.
RN	[7]			CC	EMBL; S75442; AAB20836.1; JOINED.
RP	SEQUENCE OF 591-1019 FROM N.A., AND ALTERNATIVE SPLICING.			CC	EMBL; S75444; AAB20836.1; JOINED.
RX	MEDLINE-90202932; PubMed-1690728;			CC	EMBL; S75446; AAB20836.1; JOINED.
RA	Saitta B., Stokes D.G., Vissing H., Timpl R., Chu M.-L.;			CC	EMBL; S75448; AAB20836.1; JOINED.
RA	"Alternative splicing of the human alpha 2(VI) collagen gene generates			CC	EMBL; S75450; AAB20836.1; JOINED.
RT	multiple mRNA transcripts which predict three protein variants with			CC	EMBL; S75452; AAB20836.1; JOINED.
RT	distinct carboxyl termini.";			CC	EMBL; S75454; AAB20836.1; JOINED.
RL	J. Biol. Chem. 265:6473-6480(1990).			CC	EMBL; S75456; AAB20836.1; JOINED.
RN	[8]			CC	EMBL; S75458; AAB20836.1; JOINED.
RP	VARIANT BM SER-271.			CC	EMBL; M34572; AAA35618.1; -
RX	MEDLINE-96376983; PubMed-8782832;			CC	EMBL; M34571; AAA35619.1; -
RA	Joebis G.J., Kelzers H., Vreijling J.P., de Visser M., Speer M.C.,			CC	EMBL; M34573; AAA35620.1; JOINED.
RA	Wolterman R.A., Baas F., Bohlhuis P.A.;			CC	EMBL; M34571; AAA35620.1; JOINED.
RA				CC	EMBL; M34570; AAA35621.1; -



Tue Oct 7 15:52:31 2003

```

FT DOMAIN 1 14 NONHELICAL REGION (N-TERMINAL).
FT DOMAIN 15 1040 TRIPLE-HELICAL REGION.
FT DOMAIN 1041 1049 NONHELICAL REGION (C-TERMINAL).
FT MOD_RES 95 95 HYDROXYLATION.
FT MOD_RES 107 107 HYDROXYLATION.
FT MOD_RES 119 119 HYDROXYLATION.
FT MOD_RES 938 938 HYDROXYLATION.
FT MOD_RES 950 950 HYDROXYLATION.
FT CARBOHYD 107 107 O-LINKED (GAL. . .).
FT CARBOHYD 950 950 O-LINKED (GAL. . .).
FT DISULFID 1040 1040 INTERCHAIN.
FT DISULFID 1041 1041 INTERCHAIN.
SQ SEQUENCE 1049 AA; 93651 MW; 8EEC33D1C66EC9A3 CRC64;

Query Match 13.3%; Score 77; DB 1; Length 1049;
Best Local Similarity 37.5%; Pred. No. 19;
Matches 27; Conservative 4; Mismatches 25; Indels 16; Gaps 3;

QY 13 AOGAMLAQERRVP-----RAAEVFGAQGGQGGPRG--REEAAPRGVYR-----MAVPL 56
Db 319 ANGLPGAAGERGVFGFPGPAGANGLPGEKPPGDGGPGPAGPRGVAGEPRGNLPGGPG 378

QY 57 LRRMEGAPAGPG 68
Db 379 LRGIPIGSPGGPG 390
```

Search completed: October 7, 2003, 13:26:21  
Job time : 32.7342 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:18:33 ; Search time 182.127 Seconds  
(without alignments)  
154.441 Million cell updates/sec

Title: US-09-807-512-2

Perfect score: 579

Sequence: 1 MLMQAELAFMLAQAGMLAA.....SWSAGSCPGMPLSPDQGRF 109

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	579	100.0	109	4 Q95987	O95987 homo sapien
2	264	45.6	58	4 Q95146	O95146 homo sapien
3	91.5	15.8	497	4 Q8TEP0	Q8TEP0 homo sapien
4	91	15.7	94	4 Q8IZ78	Q8IZ78 homo sapien
5	89	15.4	224	11 Q8BPA5	Q8BPA5 mus musculus
6	89	15.4	314	2 Q68475	O68475 xanthomonas
7	85.5	14.8	500	10 Q8RZJ9	Q8RZJ9 oryza sativ
8	83.5	14.4	126	11 Q9CRK3	Q9CRK3 mus musculus
9	83	14.3	312	5 Q9VRN0	Q9VRN0 drosophila
10	83	14.3	318	5 Q8MR23	Q8MR23 drosophila
11	82.5	14.2	520	11 Q8C6X3	Q8C6X3 mus musculus
12	82	14.2	180	4 Q9Y479	Q9Y479 homo sapien
13	82	14.2	210	4 Q9BU80	Q9BU80 homo sapien
14	82	14.2	210	4 Q9UJ89	Q9UJ89 homo sapien
15	82	14.2	241	10 Q9SDH5	Q9SDH5 oryza sativ
16	82	14.2	315	16 Q9RU53	Q9RU53 deinococcus

#### ALIGNMENTS

RESULT 1

O95987 PRELIMINARY; PRT; 109 AA.  
AC O95987:  
DT 01-MAY-1999 (Tremblrel. 10, Created)  
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)  
DE 01-MAY-1999 (Tremblrel. 10, Last annotation update)  
DE CTL-recognized antigen ON melanoma (CAMEL).  
GN CAMEL.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Melanoma;  
RA Aarnoudse C.A., Doel vanden P.B., Heemskerk B., Schrier P.I.;  
RT "IL-2 induced melanoma-specific CTL recognize CAMEL, an unexpected translation product of LAGE-1";  
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ012835; CAA10197.1; -  
DR EMBL; AJ012833; CAA10193.1; -  
DR EMBL; AJ012834; CAA10195.1; -  
SQ SEQUENCE 109 AA; 11689 MW; 188F85BC04C1F5F0 CRC64;

Query Match 100.0%; Score 579; DB 4; Length 109;  
Best Local Similarity 100.0%; Pred. No. 1.5e-48;  
Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLMQAELAFMLAQAGMLAAQERRVPRAAEVFGAGQQGPRGEEAPRGVRMAVPLLRM 60  
Db 1 MLMQAELAFMLAQAGMLAAQERRVPRAAEVFGAGQQGPRGEEAPRGVRMAVPLLRM 60

QY 61 EGAPAGPGRTAACFSCTSRCLSRPWRKRSWSAGSCPGMPLSPDQGRF 109  
Db 61 EGAPAGPGRTAACFSCTSRCLSRPWRKRSWSAGSCPGMPLSPDQGRF 109

RESULT 2

O95146	PRELIMINARY;	PRT;	58 AA.
ID	O95146		
AC	O95146;		
DT	01-MAY-1999 (TReMBLrel. 10, Created)		
DT	01-MAY-1999 (TReMBLrel. 10, Last sequence update)		
DT	01-MAR-2001 (TReMBLrel. 16, Last annotation update)		
DE	LAGE-2ALT protein isoform.		
GN	LAGE-2.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Wang R.-F., Johnston S.L., Zeng G., Topalian S.L.,		
RA	Schwartztruber D.J., Rosenberg S.A.;		
RT	"A Breast and Melanoma-Shared Tumor Antigen: T Cell Responses to		
RT	Antigenic Peptides Translated from Different Open Reading Frames.";		
RL	J. Immunol. 161:3596-3606(1998).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RA	Letche B.G.;		
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RA	MEDLINE=99454989; PubMed=10523621;		
RT	De Smet C., Lurquin C., Lethe B., Martelange V., Boon T.;		
RT	"DNA methylation is the primary silencing mechanism for a set of germ		
RT	line- and tumor-specific genes with a CpG-rich promoter.";		
RL	Mol. Cell. Biol. 19:7327-7335(1999).		
DR	EMBL; AF038567; AAD05203.1; -		
DR	EMBL; AJ275977; CAB76944.1; -		
SQ	SEQUENCE 58 AA; 6188 MW; ED12057564BC7EF2 CRC64;		
Query Match 45.6%; Score 264; DB 4; Length 58;			
Best Local Similarity 88.7%; Pred. No. 1.9e-18;			
Matches 55; Conservative 2; Mismatches 1; Indels 4; Gaps 1;			
QY	1 MLMQAELAFMLAQAGMLAAQERRVPRAAEVPGAQGGQGRGEEAPRGVMAVPLLRM 60		
DB	1 MLMQAELAFMLAQAGMLAAQERRVPRAAEVPGAQGGQGRGEEAPRGVMAA---RL 56		
QY	61 EG 62		
DB	57 QG 58		
RESULT 3			
Q8TEPO	PRELIMINARY;	PRT;	497 AA.
ID	Q8TEPO		
AC	Q8TEPO;		
DT	01-JUN-2002 (TReMBLrel. 21, Created)		
DT	01-JUN-2002 (TReMBLrel. 21, Last sequence update)		
DT	01-OCT-2002 (TReMBLrel. 22, Last annotation update)		
DE	FLJ00153 protein (Fragment).		
GN	FLJ00153.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Spleen;		
RA	Jikuya H., Takano J., Nomura N., Kikuno R., Nagase T., Ohara O.;		
RT	"The nucleotide sequence of a long cDNA clone isolated from human		
RT	spleen.";		
RL	Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AK074082; BAB84908.1; -		
DR	InterPro; IPR001895; RasGRF_CDC25.		
DR	InterPro; IPR000159; RA_domain.		
DR	Pfam; PF00788; RA; 1.		
DR	Pfam; PF00617; RasGEF; 1.		
DR	SMART; SM00147; RasGEF; 1.		

FT	NON_TER	1	
SQ	SEQUENCE	497 AA;	54311 MW; F78318B92826F537 CRC64;
Query Match 15.8%; Score 91.5; DB 4; Length 497;			
Best Local Similarity 30.9%; Pred. No. 0.63;			
Matches 29; Conservative 9; Mismatches 55; Indels 1; Gaps 1;			
QY	8 LAFMLAQAGMLAAQERRVPRAAEVPGAQGGQGRGEEAPRGVMAVPLLRMEGAPAG- 66		
DB	278 LAALHAQNQLTEEQSYRLSRVIEPPAAASCPSSPRIRRLISLTKRLSAKLAREKSSSPSGS 337		
QY	67 PGGRTAACFCTSRCLSRPWRKRSWSAGSCGMP 100		
DB	338 PGDPSSPTSSVSGSPSPSRSDAPAGSPASP 371		
RESULT 4			
Q8IZ78	PRELIMINARY;	PRT;	94 AA.
ID	Q8IZ78		
AC	Q8IZ78;		
DT	01-MAR-2003 (TReMBLrel. 23, Created)		
DT	01-MAR-2003 (TReMBLrel. 23, Last sequence update)		
DT	01-MAR-2003 (TReMBLrel. 23, Last annotation update)		
DE	Hypothetical protein.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Pancreas;		
RA	Strausberg R.;		
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; BC015744; AAH15744.1; -		
KW	Hypothetical protein.		
SQ	SEQUENCE 94 AA; 10329 MW; C95D33426364379D CRC64;		
Query Match 15.7%; Score 91; DB 4; Length 94;			
Best Local Similarity 33.0%; Pred. No. 0.14;			
Matches 32; Conservative 4; Mismatches 51; Indels 10; Gaps 2;			
QY	11 LMAQAGMLAAQERRVPRAAEVPGAQGGQGRGEEAPRGVMAVPLLRMEGAPAGGGR 70		
DB	1 MQTQAEALTAGMAGVATAAGAWTQPLRPVELPQTRQVRAETPRLR-----PGGH 52		
QY	71 TAACFCTSRCLSRPWRKRSWSAGSCGMPHLSPDOG 107		
DB	53 ECGRTYSPSACLSPRPWRKRSMPGPHQMP--SPTKG 87		
RESULT 5			
Q8BPA5	PRELIMINARY;	PRT;	224 AA.
ID	Q8BPA5		
AC	Q8BPA5;		
DT	01-MAR-2003 (TReMBLrel. 23, Created)		
DT	01-MAR-2003 (TReMBLrel. 23, Last sequence update)		
DT	01-MAR-2003 (TReMBLrel. 23, Last annotation update)		
DE	Hypothetical protein.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=Pituitary;		
RX	MEDLINE=22354683; PubMed=12466851;		
RA	The FANTOM Consortium,		
RT	the RIKEN Genome Exploration Research Group Phase I & II Team;		
RT	"Analysis of the mouse transcriptome based on functional annotation of		
RT	60,770 full-length cDNAs.";		
RL	Nature 420:563-573(2002).		
DR	EMBL; AK077360; BAC36767.1; -		
KW	Hypothetical protein.		

[illegible]

"Sequencing of *Drosophila melanogaster* genome." ;  
accession to the EMBL/GenBank/DBJ databases.

[illegible]

RESULT 10	PRELIMINARY;	PRT;	318 AA.
Q8MR23			
ID			
Q8MR23;			
DT 01-OCT-2002 (TREMBlrel. 22, Created)			
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)			
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)			
DE LD26105p (Fragment).			
GN CG10591.			
OS <i>Drosophila melanogaster</i> (Fruit fly).			
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC Ephydroidea; Drosophilidae; Drosophila.			
OC NCBI_Taxid=7227;			
[1]			
RP SEQUENCE FROM N.A.			

RC	STRAIN-Berkeley.
RA	Stapleton M., Brokstein P., Hong L., Aghayani A., Carlson J.,
RA	Champe M., Chavez C., Dorsett V., Dresnek D., Fagan D., Frise E.,
RA	George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA	Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA	Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA	Celniker S.;
RL	Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR	EMBL; AY122168; AAWS2680.1; -.
DR	Flybase; FBgn0035621; CGI0591.
DR	InterPro; IPR000087; Collagen.
DR	NON_TER
FT	1
SQ	SEQUENCE      318 AA;    3980 MW;    33679A1EA59E524B CRC64;

**RESULT 9**

ID	QVRVNO	PRELIMINARY;	PRT;	312 AA.
AC	QVRVNO			
DT	QVRVNO			
DT	QVRVNO			
DT	QVRVNO			
DE	CGI0591 protein.			
GN	CGI0591			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OX	Ephydroidae; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Berkley;			
RX	MEDLINE=20196006; PubMed=10731132;			
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Amanatides P.G., Scherier S.E., Li P.W., Hoskins R.A., Galie R.F.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,			
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,			
RA	Arif J.F., Aqbavani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,			
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Beeson D.I., Benos P.V., German B.P., Bhargava D., Bolshakov S.,			
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,			
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,			
RA	de Paulos B., Delcher A., Deng Z., Mays A.D., Dew I.J., Dietz S.M.,			
RA	Dodson K.C., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,			
RA	Fowler C., Gabriellista A.E., Garg N.S., Gelbart W.M., Glasser K.,			
RA	Glocke A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA	Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,			
RA	Hoskin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,			
RA	Jalali M., Kalush F., Karpen G.H., Kratz, Kennison J.A., Ketchum K.A.,			
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,			
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,			
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,			
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,			
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,			
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,			
RA	Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,			
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,			
RA	Shue B.C., Sliden-Klamis I., Simpson M., Skupski M.P., Smith T.,			
RA	Spirer E., Spradling A.C., Stapleton M., Strong R., Sun E., Wang X.,			
RA	Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,			
RA	Zhang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,			
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,			
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,			
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,			
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,			
RT	"The genome sequence of Drosophila melanogaster."			
RL	SCIENCE 287:2185-2195(2000).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,			
RA	Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,			
RA	Banzon J., An H., Baldwin D., Banzon J., Besoon K.Y., Busam D.A.,			
RA	Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,			
RA	Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,			
RA	Ferreira S., Frise E., Galie R.F., Garg N.S., George R.A.,			
RA	Gonzalez C., Hoskins R.A., Hostin D., Howard T.J.,			
RA	Thegwan C., Houck J., Hoskins R., Kruse D., Li P., Mattei B., Moshrefi A.,			
RA	McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,			
RA	Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,			
RA	Phouenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,			
RA	Stapleton M., Strong R., Swirskas R., Tector C., Tyler D.,			
RA	Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;			

```
Query Match      14.3%; Score 83; DB 5; Length 318;
Best Local Similarity 29.5%; Pred. No. 2.7;
Matches 36; Conservative 13; Mismatches 45; Indels 28; Gaps 7;

QY 8 LAFLMAQAMLAQAQRRVPRAAEVGAQGOQGRREAP---RGVRMA----- 53
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 13 LAFLVA-GLCSTAQAASLVKSCD-FPQGERGPPGPPGPNQNSGSGGYGYPPTG 70
   :|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 54 -----VPLLRMEGAPAGPGGRTAACFCTSRCLSRPWKRSWSAGSCPGMPHLS-PDQG 107
   :|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 71 PYPNLPFIQGRGLP-GPPGPGYCFPCPA-----PPPLRSFPPGVPVPFISTPGG 124
   :|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 108 RF 109
   |
Db 125 GF 126

RESULT 11
Q8C6X3 PRELIMINARY; PRT; 520 AA.
AC Q8C6X3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Procollagen (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
   60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK052963; BAC35222.1; -.
FT NON_TER 1
SQ SEQUENCE 520 AA; 52813 MW; F9803875F79244A4 CRC64;

Query Match      14.2%; Score 82.5; DB 11; Length 520;
Best Local Similarity 33.3%; Pred. No. 4.9;
Matches 25; Conservative 5; Mismatches 18; Indels 27; Gaps 3;

QY 33 GAQGOQGPGRREAPRGVRMAVPLLRMEGAPAGPGGRTAACFCTSRCLSRPWKRSWS 92
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 217 GQGGQGPKG-PEGPRG-----ETGPAGPQG-----PPGPQGPS 249
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 93 AGSCPGMPHLSPDQG 107
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 250 GLSIOGMPGMPGDKG 264
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 12
Q9Y479 PRELIMINARY; PRT; 180 AA.
AC Q9Y479;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE LAGE-1S protein (Cancer/testis antigen 2).
GN LAGE1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RX MEDLINE=99325550; PubMed=10399963;
RA Aarnoudse C.A., Van den Doel P.B., Heemskerk B., Schrier P.I.;
```

```
RT "Interleukin-2-induced, melanoma-specific T cells recognize CAMEL, an
   unexpected translation product of LAGE-1.";
RL Int. J. Cancer 82:442-448(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Aradhyia S., Bardaro T., Galgoczy P., Yamagata T., Esposito T.,
RA Patlan H., Ciccodicola A., Kenwick S., Platzer M., D'Urso M.,
RA Nelson D.L.;
RT "Multiple pathogenic and benign genomic rearrangements occur at a 35-
   kb duplication involving the NEMO and the LAGE2 genes.";
RL Hum. Mol. Genet. 0:0-0(2001).
DR EMBL; AJ012834; CAAL0194.1; -.
DR EMBL; AF277315; AAL27015.1; -.
SQ SEQUENCE 180 AA; 18236 MW; 9077FAF953543A25 CRC64;

Query Match      14.2%; Score 82; DB 4; Length 180;
Best Local Similarity 37.7%; Pred. No. 2;
Matches 23; Conservative 3; Mismatches 19; Indels 16; Gaps 2;

QY 33 GAQGOQGPGR---REEAPRGVRMAVPLLRMEGAPAGPGGRTAACFCTSRCLSRPWK 88
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 38 GATGGRGPRGAGAAASGPRG-----GAPRGPHGGAASAQDGRCPGARRPDS 85
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 89 R 89
   |
Db 86 R 86

RESULT 13
Q9BU80 PRELIMINARY; PRT; 210 AA.
AC Q9BU80;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Cancer/testis antigen 2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBDJ databases.
DR EMBL; BC002833; AA02833.1; -.
SQ SEQUENCE 210 AA; 21089 MW; 8FB5BF04FB04E8BE CRC64;

Query Match      14.2%; Score 82; DB 4; Length 210;
Best Local Similarity 37.7%; Pred. No. 2.3;
Matches 23; Conservative 3; Mismatches 19; Indels 16; Gaps 2;

QY 33 GAQGOQGPGR---REEAPRGVRMAVPLLRMEGAPAGPGGRTAACFCTSRCLSRPWK 88
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 38 GATGGRGPRGAGAAASGPRG-----GAPRGPHGGAASAQDGRCPGARRPDS 85
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 89 R 89
   |
Db 86 R 86

RESULT 14
Q9UJ89 PRELIMINARY; PRT; 210 AA.
AC Q9UJ89;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE LAGE-1L protein.
GN LAGE1L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:21:07 ; Search time 59.3291 seconds  
(without alignments)  
77.734 Million cell updates/sec

Title: US-09-807-512-2  
Perfect score: 579  
Sequence: 1 MLMAQALFLMAQGLAA.....SWSAGSCGPMPLSPDQGRF 109

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pap.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pap.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pap.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pap.\*  
5: /cgn2\_6/ptodata/1/iaa/PTUS\_COMB.pap.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	17.4	602	US-09-252-991A-24403	Sequence 24403, A
2	99	17.1	146	US-09-252-991A-19746	Sequence 19746, A
3	99	17.1	359	US-08-586-165-4	Sequence 4, Appli
4	94.5	16.3	431	US-09-252-991A-25721	Sequence 25721, A
5	94.5	16.3	510	US-09-252-991A-25466	Sequence 25466, A
6	94	16.2	345	US-09-252-991A-32325	Sequence 32325, A
7	94	16.2	699	US-09-252-991A-26231	Sequence 26231, A
8	92.5	16.0	849	US-09-252-991A-21865	Sequence 21865, A
9	92	15.9	343	US-09-252-991A-24630	Sequence 24630, A
10	91.5	15.8	324	US-09-252-991A-24664	Sequence 24664, A
11	91	15.7	1694	US-08-494-168-2	Sequence 2, Appli
12	90.5	15.6	250	US-09-252-991A-30771	Sequence 30771, A
13	90	15.5	254	US-09-252-991A-26135	Sequence 26135, A
14	90	15.5	301	US-09-252-991A-30700	Sequence 30700, A
15	89.5	15.5	312	US-09-252-991A-28271	Sequence 28271, A
16	89.5	15.5	321	US-09-252-991A-23173	Sequence 23173, A
17	89.5	15.5	538	US-09-252-991A-24079	Sequence 24079, A
18	88.5	15.3	440	US-09-252-991A-30167	Sequence 30167, A
19	88	15.2	147	US-09-252-991A-26570	Sequence 26570, A
20	88	15.2	416	US-09-252-991A-30923	Sequence 30923, A
21	87.5	15.1	219	US-09-252-991A-23215	Sequence 23215, A
22	87.5	15.1	259	US-09-252-991A-24558	Sequence 24558, A
23	87.5	15.1	457	US-09-252-991A-25131	Sequence 25131, A
24	87.5	15.1	494	US-09-252-991A-22209	Sequence 22209, A
25	87	15.0	245	US-09-252-991A-19505	Sequence 19505, A
26	87	15.0	293	US-09-252-991A-23363	Sequence 23363, A
27	86.5	14.9	343	US-09-252-991A-23162	Sequence 23162, A

ALIGNMENTS

RESULT 1

US-09-252-991A-24403  
; Sequence 24403, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 24403  
; LENGTH: 602  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-24403

Query Match 17.4%; Score 101; DB 4; Length 602;  
Best Local Similarity 34.1%; Pred. No. 0.0065;  
Matches 29; Conservative 6; Mismatches 40; Indels 10; Gaps 3;  
QY 23 RRVPRAAEYVGAQGQGRGEEAPRGVMAVPLLRMEGAPAGPGGRTAA--CF----- 75  
Db 29 RTAHRRESAPARRRASPRRRPLPRRAGRSVPGCRRNAASPPGVSVPRAASRCVRRGR 88  
QY 76 -SCTSR--CLSRPWRKRSWSAGSCP 97  
Db 89 RACARRGCCSRAPWPRRSATHPCP 113

RESULT 2

US-09-252-991A-19746  
; Sequence 19746, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27

28 86 14.9 175 4 US-09-252-991A-16749 Sequence 16749, A  
29 86 14.9 179 4 US-09-252-991A-25925 Sequence 25925, A  
30 86 14.9 611 4 US-09-252-991A-17597 Sequence 17597, A  
31 86 14.9 710 4 US-09-252-991A-24946 Sequence 24946, A  
32 85 14.7 863 4 US-09-252-991A-26099 Sequence 26099, A  
33 84.5 14.6 145 4 US-09-252-991A-23153 Sequence 23153, A  
34 84.5 14.6 166 4 US-09-252-991A-18554 Sequence 18554, A  
35 84.5 14.6 171 4 US-09-252-991A-29469 Sequence 29469, A  
36 84.5 14.6 321 4 US-09-252-991A-17328 Sequence 17328, A  
37 84 14.5 169 4 US-09-252-991A-22543 Sequence 22543, A  
38 84 14.5 311 4 US-09-252-991A-22295 Sequence 22295, A  
39 84 14.5 886 4 US-09-252-991A-26999 Sequence 26999, A  
40 83.5 14.4 688 4 US-09-252-991A-21780 Sequence 21780, A  
41 83 14.3 187 4 US-09-252-991A-29271 Sequence 29271, A  
42 83 14.3 246 4 US-09-252-991A-20289 Sequence 20289, A  
43 83 14.3 412 4 US-09-252-991A-27252 Sequence 27252, A  
44 83 14.3 511 4 US-09-252-991A-20054 Sequence 20054, A  
45 83 14.3 587 4 US-09-252-991A-25368 Sequence 25368, A

```

; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19746
; LENGTH: 146
; TYPE: PRN
; ORGANISM: Pseudomonas aerug
US-09-252-991A-19746

```

Query Match	17.18	Score 99	DB 4	Length 146
Best Local Similarity	34.88	Pred. No. 0.0021		
Matches	32	Conservative	36	Indels 20
		Mismatches	36	Gaps 5

  

QY	24	RVPAAEYVPAQGO-QGPRGREAPRGVNRVAVLLRRMEGAPAGCGGRTACFCTSRCL	82
Db	12	RSPPAPRYGARRCPGSPGSRGGPRERR-----RRSRGRPR-----RNR-----RCRTKSA	58
QY	83	SRRPWK-----RSWSAGSCPGMPLHSPDQGR	108
Db	59	CRPPNKSAPGPGRGWPAACCRPSPAPCAGR	90

RESULT 3  
US-08-586-165-4

```

US-08-366-103-4
; Sequence 4, Application US/08586165
; Patent No. 6054298
; GENERAL INFORMATION:
; APPLICANT: Laufer, Edward M.
; APPLICANT: Orozco, Olivia E.
; APPLICANT: Tabin, Clifford J.
; TITLE OF INVENTION: Fringe Proteins and Pattern Formation
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
;

```

ATTORNEY/AGENT INFORMATION:  
NAME: Granahan, Patricia  
REGISTRATION NUMBER: 327227  
REFERENCE/DOCKET NUMBER: H095-05  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 861-6240  
TELEFAX: (617) 861-9340  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 359 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1118  
US-08-586-165-4

	Query Match	17.1%	Score 99;	DB 3;	Length 359;	
	Best Local Similarity	35.2%;	Pred. No. 0.0059;			
	Matches	38; Conservative	6; Mismatches	42; Indels	22; Gaps	6;
QY	20 AQERRVRAAEVPGAQGQQGPRG-----REAPRGVRMAVPLLRMEGPAG--P	67				
DH	:					
	37 AAARPPCPAC-PPRSGRRGRGCATCGTAAARE--PRGAAGASPVARGLRCGSANGLP	93				

```

QY 58 GGRTAACFSCTSRCLSRP--WKRSWSAGSCPGMP-----HLSPDGR 108
      | : ||| ||| : ||| ||| | ||| |
DB 94 AARPRAWKTSLLRRKRGSI TRRGWSCCKPGSPGREDRHSYSQTR 141

```

RESULT 4  
US-09-252-991A-25721  
Sequence 25721, Application US/09252991A  
Patent NO. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELAT  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 25721  
LENGTH: 431  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-25721

	Query Match	16.3%;	Score 94.5;	DB 4;	Length 431;
	Best Local Similarity	29.4%;	Pred. No. 0.022;		
Matches	Conservative	7;	Mismatches	44;	Indels 45; Gaps 8;
QY	15	GAMLAQAQERRYPRAAEVPG--AQCQGCPGRGEEAPGVMAVPLLRRMEGAPAGPGR--	-	70	
		:                                 :			
Db	175	GAGRPGERGRRGAAPFAFLRVQGG-AGRRRKRQRLLARAGRCGGKSACAGGRP		233	
		:                                 :			
QY	71	-----TAACFCSTSRCLS-----RRP-----			
Db	234	RRLDGRGAC--CRARCSNAMPPTACRRRPWCWPGRSPGSSRATRDSLWKWDRWSRGAC		291	
QY	97	----PGMPHLSPDQGR	108		
Db	292	STSSPSWPDPSGSGRR	307		

```

RESULT 5
US-09-252-991A-25466
; Sequence 25466, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELAT
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25466
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25466

```

	Query Match	16.3%	Score 94.5	DB 4	Length 510
	Best Local Similarity	31.1%	Pred. No. 0.027	53	Indels 21
	Matches 37	Conservative	8	Mismatches	Gaps 4
QY	11	LMAQAML-----AAQERR-----	VPRAAEVPGAQQQGP-RGRREAPGRVMAY	54	
ph	42	LRAAGGVPRPAAHRRRFGHVRVAPGFGFPRPVAQAPGDRRGRTGCTGRAGQAQSAHGR	101		



[illegible]

```

RESULT 6
US-09-252-991A-32325
; Sequence 32325, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32325
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32325

```

```

Query Match      16.2%; Score 94; DB 4; Length 345;
Best Local Similarity 39.4%; Pred. No. 0.019;
Matches 28; Conservative 4; Mismatches 25; Indels 14; Gaps 4;

QY 26 PRAEVPGAQGQ-----QGPRGR-----EAPRGVRMAVPLLRMEGAPAGPGGRTAACF 75
Db 221 PVGGKIPFAGRGCAQSPARRALAPERAGRSARAG---HRRPGDLRGGGAARC- 276

QY 76 SCTSRCLSRRP 86
Db 277 RCRHRTAARRP 287

```

RESULT 7  
 US-09-252-991A-26231  
 ; Sequence 26231, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; CURRENT FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 26231  
 ; LENGTH: 699  
 ; TYPE: PRT  
 ; ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-26231

[illegible]

```

Db      441  GTDDRGGAGPGVCAPTAADPTGTAC--RPWRTTASGAORRASSDGAERPCP-L 457

QY      100  PHL-----SPDQGR 108

Db      498  PHCPRRARRGCGEEPACGRDPQGR 521

```

```

RESULT 8
US-09-252-991A-21865
; Sequence 21865, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21865
; LENGTH: 849
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21865

```

	Query Match	16.0%;	Score 92.5;	DB 4;	Length 849;
	Best Local Similarity	33.7%;	Pred. No. 0.08;		
	Matches 3;	Conservative	5;	Mismatches 45;	Indels 13; Gaps 5;
Qy	18	LAQERVPRAAEVPVGAQGOQGPGRREAPRGVYMAVPL--LRMEGAPA---GPGRTA	72		
Db	148	IRATRRFAARSRSAAGAAPPSPGGRRARARCPAGVPRRGIRSRPGPAPGAPAGR-A	206		
Qy	73	ACFSCTSRCLSRPWRKWSAGSCPGMPLSPDQG	107		
Db	207	GC-SCRSR-----WGRRWPAGDPRAIPAPAPAGG	234		

```

RESULT 9
US-09-252-991A-24630
; Sequence 24630, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24630
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24630

```

```

Query Match      15.9%; Score 92; DB 4; Length 343;
Best Local Similarity 31.2%; Pred. No. 0.032;
Matches 44; Conservative 6; Mismatches 37; Indels 54; Gaps 7;

Qy 11 LMAQGMALAAQE-----RRV--PRAAEVVP-----GAQOQQQPRGR 43
      | : | : | : | : | : | : | : | : | : | : | : | : |
Db 99 LVAGRMARGARESLREFLRHSVGIPIRRRVAPRGAAVPAQLRRGPGCGDRPEPGR 158
      | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 44 EAPRGVYRMVAPLL-----RRMEGAPAGPGG--RTAAACFCTSR 81

```

Db	159	PPAGEPRRAALPLAQLPGAAGRRAAPVSPGRFRPRVPAGRGLARDAA---STGGC	215
QY	82	LSRRPWKRS--WSAGSCGMP	100
Db	216	ARRERSARSPPAHAGRTGTP	236

RESULT 10  
US-09-252-991A-24664  
; Sequence 24664, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 24664  
; LENGTH: 324  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-24664

```

RESULT 11
US-08-494-168-2
; Sequence 2, Application US/08494168
; Patent No. 5731192
; GENERAL INFORMATION:
; APPLICANT: Reeders, Stephen T.
; APPLICANT: Zhou, Jing
; TITLE OF INVENTION: Collagen COL4A6: Gene, Protein and Method
; TITLE OF INVENTION: of Detecting Collagen Deficiency
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/494,168
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/112,465
; FILING DATE: 27-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D.

```

```

/      REGISTRATION NUMBER: 28,665
/
/      REFERENCE/DOCKET NUMBER: 403997/104/BABR
/
/      TELECOMMUNICATION INFORMATION:
/
/      TELEPHONE: (202)672-5300
/
/      TELEFAX: (202)672-5399
/
/      TELEX: 904136
/
/      INFORMATION FOR SEQ ID NO: 2:
/
/      SEQUENCE CHARACTERISTICS:
/
/      LENGTH: 1694 amino acids
/
/      TYPE: amino acid
/
/      TOPOLOGY: linear
/
/      MOLECULE TYPE: protein
/
US-08-494-168-2

```

RESULT 12  
US-09-252-991A-30771  
; Sequence 30771, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 30771  
; LENGTH: 250  
; TYPE: prt  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-30771

```

RESULT 13
US-09-252-991A-26135
; Sequence 26135, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18

```

; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26135
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26135

Query Match 15.5%; Score 90; DB 4; Length 254;
Best Local Similarity 26.9%; Pred. No. 0.037;
Matches 35; Conservative 6; Mismatches 53; Indels 36; Gaps 3;

QY 15 GAMLAAQERRVPRAAEVPGAGQGQGRGEEAPRGVPM---AVPLLRMEGAPAGPG--- 68
Db 49 GVCRAAPGHRAAAARVPLGHTSGGDFRDRPRGARIPATGLPHRRRSPSPAAPHGK 108
QY 69 -----GRTAACFCTSRCL-----SRRPWKRSWSAGSCFG 98
Db 109 TPSNAGAAASRPAGRAVPGSRRRRCPLPARAPGPRHPPARGMARRPAAEGRPAPAGAG 168
QY 99 MHLSPDQGR 108
Db 169 RPSRRPPAGR 178

RESULT 14
US-09-252-991A-30700
; Sequence 30700, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30700
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30700

Query Match 15.5%; Score 90; DB 4; Length 301;
Best Local Similarity 27.0%; Pred. No. 0.045;
Matches 40; Conservative 11; Mismatches 43; Indels 54; Gaps 6;

QY 14 QGAMLAQERRVPRAAEVPGAQ---GQGGPRGEEAPRGVPMVPLLRMEGAPAGPGGR 70
Db 70 QCADRRAMHRLPRAGTDPQGRYRQHRSCRSA-RYRRQLPWLRLHALGSPAHPRAA 128
QY 71 TAACES-----CTSRCL-----RRPWK----- 88
Db 129 TGPYAGAPLPGAGTQRSALAIQVLPAGLPHRTGRODPRPAPRWRTGRGGGAAGGVR 188
QY 89 -----RSWSAGSCGMPHL---SPDQGR 108
Db 189 HARAARSVARTACGPGPPAANWTPATGR 216

RESULT 15
US-09-252-991A-28271
; Sequence 28271, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28271
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28271

Query Match 15.5%; Score 89.5; DB 4; Length 312;
Best Local Similarity 29.3%; Pred. No. 0.053;
Matches 41; Conservative 4; Mismatches 42; Indels 53; Gaps 8;

QY 13 AOGAMLAQERRVPRAAEVPGAGQGQGRGEEAPRGVPMVPLLRMEGAPAGPG--- 68
Db 44 AOGARRTGARR--PRAA-APG-QGHFPGCGADLDGRRQR---LHRVALGPGPPGPGPA 96
QY 69 -----GRTAACFCTSR-----CLSRRPWKR-----SWS 92
Db 97 AGGAPAPGRRRRCHPPGDRRRRAGAGTDPALPRRPAEPDPPDLLRRFPQRAARRLPGRF 156
QY 93 AGSCFGM-----PHLSPDQ 106
Db 157 AASLFCGRQRPAAADPHLQPDOR 176

Search completed: October 7, 2003, 13:34:04
Job time : 60.3291 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:31:20 ; Search time 118.658 Seconds  
(without alignments)  
145.335 Million cell updates/sec

Title: US-09-807-512-2

Perfect score: 579

Sequence: 1 MLMAQELAFMLAQAMLA.....SWSAGSCFPMHLSPDQGRF 109

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues

Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications\_AA:\*
- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
  - 2: /cgn2\_6/ptodata/1/pubpaa/FCT\_NEW\_PUB.pep.\*
  - 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/ptodata/1/pubpaa/FCTUS\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
  - 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
  - 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
  - 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
  - 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
  - 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
  - 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
  - 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
  - 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
  - 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	94	16.2	594	12	US-10-259-165-334
2	92.5	16.0	522	15	US-10-156-761-8465
3	91.5	15.8	699	15	US-10-233-131-34
4	91.5	15.8	710	14	US-10-118-328-2
5	91.5	15.8	710	15	US-10-060-990-3
6	91.5	15.8	715	15	US-10-233-131-33
7	82	14.2	180	15	US-10-146-473-69
8	82	14.2	210	15	US-10-157-031-88
9	81.5	14.1	709	14	US-10-118-328-4
10	81.5	14.1	3298	15	US-10-149-819-21
11	81.5	14.1	3312	15	US-10-225-567A-656
12	81	14.0	1298	10	US-09-825-288A-2
13	80.5	13.9	144	15	US-10-156-761-9586
14	80.5	13.9	297	10	US-09-738-626-6257
15	79.5	13.7	337	10	US-09-814-777A-126

16	79.5	13.7	462	10	US-09-814-777A-99
17	78.5	13.6	504	15	US-10-156-761-12417
18	78	13.5	1466	12	US-10-301-822-33
19	78	13.5	1466	15	US-10-177-293-68
20	77	13.3	180	12	US-09-849-602-30
21	77	13.3	180	15	US-10-207-655-71
22	77	13.3	397	9	US-09-821-883-27
23	77	13.3	519	15	US-10-114-170-265
24	77	13.3	1366	12	US-10-301-822-31
25	77	13.3	1366	15	US-10-171-311-38
26	77	13.3	1366	15	US-10-058-124-19
27	76	13.1	325	10	US-09-927-112-10
28	76	13.1	714	15	US-10-233-885-44
29	76	13.1	714	15	US-10-231-581-44
30	76	13.1	1014	12	US-10-194-441A-1
31	76	13.1	1207	10	US-09-927-112-2
32	76	13.1	1418	15	US-10-058-124-20
33	75.5	13.0	283	9	US-09-925-301-979
34	75.5	13.0	285	15	US-10-106-698-6330
35	75.5	13.0	666	15	US-10-225-630-10
36	75	13.0	1078	15	US-10-058-124-21
37	75	13.0	1497	15	US-10-157-031-128
38	74.5	12.9	231	10	US-09-893-737-154
39	74.5	12.9	1667	15	US-10-156-761-9873
40	73.5	12.7	237	9	US-09-220-920-32
41	73.5	12.7	237	9	US-09-804-615-4
42	73.5	12.7	238	10	US-09-813-398-40
43	73.5	12.7	912	14	US-10-047-542-75
44	73	12.6	187	11	US-09-975-719-287
45	73	12.6	430	9	US-09-925-302-518

ALIGNMENTS

RESULT 1

- US-10-259-165-334  
; Sequence 334, Application US/10259165  
; Publication No. US20030135888A1  
; GENERAL INFORMATION:  
; APPLICANT: Zhu, Tong  
; APPLICANT: Wang, Xun  
; APPLICANT: Chang, Hur-song  
; APPLICANT: Briggs, Steven P.  
; APPLICANT: Cooper, Bret  
; APPLICANT: Glazebrook, Jane  
; APPLICANT: Goff, Stephen A.  
; APPLICANT: Katagiri, Fumiyaki  
; APPLICANT: Kreps, Joel  
; APPLICANT: Moughamer, Todd  
; APPLICANT: Provart, Nicholas  
; APPLICANT: Ricke, Darrell  
; TITLE OF INVENTION: GENES THAT ARE MODULATED BY POSTTRANSCRIPTIONAL GENE SILENCING  
; FILE REFERENCE: 70030-NP  
; CURRENT APPLICATION NUMBER: US/10/259,165  
; CURRENT FILING DATE: 2002-09-26  
; PRIOR APPLICATION NUMBER: US 60/370,620  
; PRIOR FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: US 60/368,327  
; PRIOR FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: US 60/325,277  
; PRIOR FILING DATE: 2001-09-26  
; NUMBER OF SEQ ID NOS: 782  
; SOFTWARE: PatentList.pl version 3.0.4 (C) 2001 Syngenta  
; SEQ ID NO 334  
; LENGTH: 594  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
US-10-259-165-334

Query Match 16.2%; Score 94; DB 12; Length 594;  
Best Local Similarity 33.0%; Pred. No. 0.72;  
Matches 29; Conservative 8; Mismatches 37; Indels 14; Gaps 5;

```

; CURRENT APPLICATION NUMBER: US/10/233,131
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 09/808,701
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 34
; LENGTH: 699
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-233-131-34

Query Match 15.8%; Score 91.5; DB 15; Length 699;
Best Local Similarity 30.9%; Pred. No. 1.4; Indels 1; Gaps 1;
Matches 29; Conservative 9; Mismatches 55;

QY 8 LAFLMAQGAAMLAQAQERRVPRAAEVPGAQGGQGGPRGEEAPRGVMAVPLLRMEGAPAG- 66
DB 486 LAALHAQNQLTTEQSYRLSRVIEPPAAACSPSSPRIRRRISLTIKRLSAKLAREKSSSPSGS 545

QY 67 PGGRRTAACFCTSRCLSRRLRPWKRSWSAGSCPGMP 100
DB 546 PGDPSSPTSSVSPGSPSPSRSDAPAGSPASP 579

RESULT 4
US-10-118-328-2
; Sequence 2, Application US/10118328
; Publication No. US20020169289A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CLO01220
; CURRENT APPLICATION NUMBER: US/10/118,328
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/282,460
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 710
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-118-328-2

Query Match 15.8%; Score 91.5; DB 14; Length 710;
Best Local Similarity 30.9%; Pred. No. 1.5; Indels 1; Gaps 1;
Matches 29; Conservative 9; Mismatches 55;

QY 8 LAFLMAQGAAMLAQAQERRVPRAAEVPGAQGGQGGPRGEEAPRGVMAVPLLRMEGAPAG- 66
DB 481 LAALHAQNQLTTEQSYRLSRVIEPPAAACSPSSPRIRRRISLTIKRLSAKLAREKSSSPSGS 540

QY 67 PGGRRTAACFCTSRCLSRRLRPWKRSWSAGSCPGMP 100
DB 541 PGDPSSPTSSVSPGSPSPSRSDAPAGSPASP 574

RESULT 5
US-10-060-990-3
; Sequence 3, Application US/10060990
; Publication No. US20030032159A1
; GENERAL INFORMATION:
; APPLICANT: Gu, Yizhong
; APPLICANT: Nguyen, Cung-Tuong
; TITLE OF INVENTION: HUMAN RALGDS-LIKE PROTEIN 3
; FILE REFERENCE: PB0176

; CURRENT APPLICATION NUMBER: US/10/233,131
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 09/808,701
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 34
; LENGTH: 699
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-233-131-34

Query Match 15.8%; Score 91.5; DB 15; Length 699;
Best Local Similarity 30.9%; Pred. No. 1.4; Indels 1; Gaps 1;
Matches 29; Conservative 9; Mismatches 55;

QY 8 LAFLMAQGAAMLAQAQERRVPRAAEVPGAQGGQGGPRGEEAPRGVMAVPLLRMEGAPAG- 66
DB 486 LAALHAQNQLTTEQSYRLSRVIEPPAAACSPSSPRIRRRISLTIKRLSAKLAREKSSSPSGS 545

QY 67 PGGRRTAACFCTSRCLSRRLRPWKRSWSAGSCPGMP 100
DB 546 PGDPSSPTSSVSPGSPSPSRSDAPAGSPASP 579

RESULT 4
US-10-118-328-2
; Sequence 2, Application US/10118328
; Publication No. US20020169289A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CLO01220
; CURRENT APPLICATION NUMBER: US/10/118,328
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/282,460
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 710
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-118-328-2

Query Match 15.8%; Score 91.5; DB 14; Length 710;
Best Local Similarity 30.9%; Pred. No. 1.5; Indels 1; Gaps 1;
Matches 29; Conservative 9; Mismatches 55;

QY 8 LAFLMAQGAAMLAQAQERRVPRAAEVPGAQGGQGGPRGEEAPRGVMAVPLLRMEGAPAG- 66
DB 481 LAALHAQNQLTTEQSYRLSRVIEPPAAACSPSSPRIRRRISLTIKRLSAKLAREKSSSPSGS 540

QY 67 PGGRRTAACFCTSRCLSRRLRPWKRSWSAGSCPGMP 100
DB 541 PGDPSSPTSSVSPGSPSPSRSDAPAGSPASP 574

RESULT 5
US-10-060-990-3
; Sequence 3, Application US/10060990
; Publication No. US20030032159A1
; GENERAL INFORMATION:
; APPLICANT: Gu, Yizhong
; APPLICANT: Nguyen, Cung-Tuong
; TITLE OF INVENTION: HUMAN RALGDS-LIKE PROTEIN 3
; FILE REFERENCE: PB0176

```

```

RESULT 6
US-10-233-131-33
; Sequence 33, Application US/10233131
; Publication No. US20030096279A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Xue, Aidong J.
; APPLICANT: Ren, Feiyan
; APPLICANT: Wang, Dunrui
; APPLICANT: Chen, Rui-hong
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030096279A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 790CIP2D DIVA
; CURRENT APPLICATION NUMBER: US/10/233,131
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 09/808,701
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 33

```

	Query Match	14.2%	Score 82;	DB 15;	Length 180;
	Best Local Similarity	37.7%	Prod. No. 2.9;		
	Matches 23;	Conservative	3;	Mismatches 19;	Indels 16; Gaps 2
Qy	33	GAQQQQPRG----	REAPRCGRVMAVPLRRMEGAPAGPGGRTACF	SCTSRCLSRRPWK	88
Db	38	GATGGRPGRGAARASGPR-----	GAPRGPHGGAASAOBGRPCGAGARRPDS	85	

```

> RESULT 8
> US-10-157-031-88
> Sequence 88, Application US/10157031
> Publication No. US20030108890A1
> GENERAL INFORMATION:
> APPLICANT: Baranova, A. V.
> APPLICANT: Yankovsky, N. K.
> APPLICANT: Kozlov, A. P.
> APPLICANT: Lobashev, A. V.
> APPLICANT: Krukovskaya, L. L.
> TITLE OF INVENTION: In silico screen
> FILE REFERENCE: 2760-103
> CURRENT APPLICATION NUMBER: US/10/157-031
> CURRENT FILING DATE: 2002-05-30
> NUMBER OF SEQ ID NOS: 415
> SOFTWARE: PatentIn version 3.1
> SEQ ID NO 88

```

```
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-157-031-88

Query Match      14.2%; Score 82; DB 15; Length 210;
Best Local Similarity 37.7%; Pred. No. 3.4;
Matches 23; Conservative 3; Mismatches 19; Indels 16; Gaps 2;

QY 33 GAQGOQPRG-----REEAPRGVMAVPLLRMEGAPAGPGGRTACFSCTSRCLSRPWK 88
   |||:||||| | ||| ||||| | | | | | | | | | | | | | | | | | | |
Db 38 GATGGRGPRGAGAAASGPRG-----GAPRGPHGGAASQAQDGRCPGARRPDS 85

QY 89 R 89
   |
Db 86 R 86

RESULT 9
US-10-118-328-4
; Sequence 4, Application US/10118328
; Publication No. US20020169289A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: C0001220
; CURRENT APPLICATION NUMBER: US/10/118,328
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/282,460
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 709
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-118-328-4

Query Match      14.1%; Score 81.5; DB 14; Length 709;
Best Local Similarity 28.9%; Pred. No. 12;
Matches 28; Conservative 6; Mismatches 46; Indels 17; Gaps 1;

QY 8 LAFLMAQAMLAQERRVPRAAEVPGAQGOQGRGEEAPRGVMAVPLLRMEGAPAGP 67
   ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 479 LAALRAQQLSEEQSYRVSRVIEPPAASCPSPIRRRISITKFLSAKLSREKNSPPGS 538

QY 68 GGRTAACFSCTSRCLSRPWKRSWSAGSCPGMPLSP 104
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 539 PGDPSSPTSSVS-----PGSPPSP 558

RESULT 10
US-10-149-819-21
; Sequence 21, Application US/10149819
; Publication No. US20030044913A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: TANG, Y. Tom
; APPLICANT: PATTERSON, Chandra
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: SHAH, Purvi
; APPLICANT: LAL, Preeti
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: BURFORD, Neil
; TITLE OF INVENTION: EXTRACELLULAR MATRIX AND CELL ADHESION MOLECULES
; FILE REFERENCE: PF-0760 PCT
; CURRENT APPLICATION NUMBER: US/10/149,819
; CURRENT FILING DATE: 2002-06-10

; PRIOR APPLICATION NUMBER: 60/172,852; 60/172,354
; PRIOR FILING DATE: 1999-12-10; 1999-12-16
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 21
; LENGTH: 3298
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No. US20030044913A1 2847752CD1
US-10-149-819-21

Query Match      14.1%; Score 81.5; DB 15; Length 3298;
Best Local Similarity 34.4%; Pred. No. 56;
Matches 31; Conservative 6; Mismatches 38; Indels 15; Gaps 5;

QY 16 AMLAAQERRVPRAAEVPGAQGOQGRGEEAPRGVMA-----VPLLRMEGAPAGPGR 70
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 219 ATTGAERTAPRRNCLPGASG--SGPE-LDSAPTARTAPASGSAPRESRTAPEPAKRM 276

QY 71 TAACFCTSRCLSRPWKRSWSAGSCPGMP 100
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 277 SRGLFRC--RFLPQRPGRP-----PGLP 298

RESULT 11
US-10-225-567A-656
; Sequence 656, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glenn C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPT
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 656
; LENGTH: 3312
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-225-567A-656

Query Match      14.1%; Score 81.5; DB 15; Length 3312;
Best Local Similarity 34.4%; Pred. No. 56;
Matches 31; Conservative 6; Mismatches 38; Indels 15; Gaps 5;

QY 16 AMLAAQERRVPRAAEVPGAQGOQGRGEEAPRGVMA-----VPLLRMEGAPAGPGR 70
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 220 ATTGAERTAPRRNCLPGASG--SGPE-LDSAPTARTAPASGSAPRESRTAPEPAKRM 277

QY 71 TAACFCTSRCLSRPWKRSWSAGSCPGMP 100
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 278 SRGLFRC--RFLPQRPGRP-----PGLP 299

RESULT 12
US-09-825-288A-2
; Sequence 2, Application US/09825288A
; Publication No. US20020192822A1
; GENERAL INFORMATION:
; APPLICANT: LEOPARDI, ROSARIO
; APPLICANT: ROITZMAN, BERNARD
; TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICP4 IS AN INHIBITOR OF APOPTOSIS
; FILE REFERENCE: ARCD:317USC1
; CURRENT APPLICATION NUMBER: US/09/825,288A
; CURRENT FILING DATE: 2001-04-02
```



;; PRIOR APPLICATION NUMBER: 09/259,821  
;; PRIOR FILING DATE: 1999-03-01  
;; PRIOR APPLICATION NUMBER: 08/690,473  
;; PRIOR FILING DATE: 1996-07-26  
;; NUMBER OF SEQ ID NOS: 2  
;; SOFTWARE: PatentIn ver. 2.1  
;; SEQ ID NO 2  
;; LENGTH: 1298  
;; TYPE: PRT  
;; ORGANISM: HERPES VIRUS, TYPE 1  
US-09-825-288A-2

Query Match 14.0%; Score 81; DB 10; Length 1298;  
Best Local Similarity 29.2%; Pred. No. 25;  
Matches 35; Conservative 11; Mismatches 46; Indels 28; Gaps 6;  
QY 4 AQEALAFMAQAMLAQAQERRVPRAAEVPGAGQQGPGRE---EAPRGVMAVPLLRM 60  
Db 710 AADALAAAASAPREGRRKSPGPAPPEGGGPRPKTKSGADAP-GSDARAPLPAPA 768  
QY 61 -----EGAPAGPGGRTAACFCTSR--CLSRP-----WKRSWSAGSCPGMPH 101  
Db 769 PPSTPPGPEPAPAQAAPRAAAQAARPRVAVSRRAECPDLGGWR-----QPPGPSH 823

RESULT 13  
US-10-156-761-9586  
;; Sequence 9586, Application US/10156761  
;; Publication No. US20030119018A1  
;; GENERAL INFORMATION:  
;; APPLICANT: OMURA, SATOSHI  
;; APPLICANT: IKEDA, HARUO  
;; APPLICANT: ISHIKAWA, JUN  
;; APPLICANT: HORIKAWA, HIROSHI  
;; APPLICANT: SHIBA, TADAYOSHI  
;; APPLICANT: SAKAKI, YOSHIYUKI  
;; APPLICANT: HATTORI, MASAHIRA  
;; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
;; FILE REFERENCE: 249-262  
;; CURRENT APPLICATION NUMBER: US/10/156,761  
;; CURRENT FILING DATE: 2002-05-29  
;; PRIOR APPLICATION NUMBER: JP 2001-204089  
;; PRIOR FILING DATE: 2001-05-30  
;; PRIOR APPLICATION NUMBER: JP 2001-272697  
;; PRIOR FILING DATE: 2001-08-02  
;; NUMBER OF SEQ ID NOS: 15109  
;; SEQ ID NO 9586  
;; LENGTH: 144  
;; TYPE: PRT  
;; ORGANISM: Streptomyces avermitilis  
US-10-156-761-9586

Query Match 13.9%; Score 80.5; DB 15; Length 144;  
Best Local Similarity 28.8%; Pred. No. 3.2;  
Matches 34; Conservative 8; Mismatches 35; Indels 41; Gaps 6;  
QY 14 QGAMLAQAQERRVPRAAEV-----PGAQGOQGPGRGEEAPRGVMAVPLLR 58  
Db 38 QSARVRLSQRVKQSEELLTDLEALIGHRTDCPGQGVPGD-----AAPCQR 85  
QY 59 RMEGAPA-----GPG--GRTAACFCTSRCLSRPWRKRSWSAGSCPG---MPHLSP 104  
Db 86 EPQGDATDRVGARGPGAAGTASDCLCCALTALLRR--HRSGTGGRDPGDKRSPSPSP 141

RESULT 14  
US-09-738-626-6257  
;; Sequence 6257, Application US/09738626  
;; Publication No. US20020197605A1  
;; GENERAL INFORMATION:  
;; APPLICANT: NAKAGAWA, SATOSHI  
;; APPLICANT: MIZOGUCHI, HIROSHI  
;; APPLICANT: ANDO, SEIKO

;; APPLICANT: HAYASHI, MIKIRO  
;; APPLICANT: OCHIALI, KEIKO  
;; APPLICANT: YOKOI, HARUHIKO  
;; APPLICANT: TATEISHI, NAKO  
;; APPLICANT: SENOH, AKIHIRO  
;; APPLICANT: IKEDA, MASATO  
;; APPLICANT: OZAKI, AKIO  
;; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
;; FILE REFERENCE: 249-125  
;; CURRENT APPLICATION NUMBER: US/09/738,626  
;; CURRENT FILING DATE: 2000-12-18  
;; PRIOR APPLICATION NUMBER: JP 99/377484  
;; PRIOR FILING DATE: 1999-12-16  
;; PRIOR APPLICATION NUMBER: JP 00/159162  
;; PRIOR FILING DATE: 2000-04-07  
;; PRIOR APPLICATION NUMBER: JP 00/280988  
;; PRIOR FILING DATE: 2000-08-03  
;; NUMBER OF SEQ ID NOS: 7059  
;; SOFTWARE: PatentIn ver. 3.0  
;; SEQ ID NO 6257  
;; LENGTH: 297  
;; TYPE: PRT  
;; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-6257

Query Match 13.9%; Score 80.5; DB 10; Length 297;  
Best Local Similarity 34.2%; Pred. No. 6.5;  
Matches 26; Conservative 1; Mismatches 30; Indels 19; Gaps 3;  
QY 32 PGAQGOQGPGRGEEAPRGVMAVPLLRMEGAP-----AGPGRTAACFCTSRCLSR 84  
Db 155 PGTIGTIGPSGRTGKP-----IPRESRAVITPSAAANPNAEPHNTTACACCTELCGSR 208  
QY 85 RPWKRWSWSAGSCPGMP 100  
Db 209 R-----SVSLVPGPP 218

RESULT 15  
US-09-814-777A-126  
;; Sequence 126, Application US/09814777A  
;; Patent No. US20020142415A1  
;; GENERAL INFORMATION:  
;; APPLICANT: KOOPMAN, Peter Anthony  
;; APPLICANT: MUSCAT, George Eugene Orlando  
;; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND POLYNUCLEOTIDES AND METHODS OF USING THE  
;; FILE REFERENCE: 21415-0003  
;; CURRENT APPLICATION NUMBER: US/09/814,777A  
;; CURRENT FILING DATE: 2001-03-23  
;; PRIOR APPLICATION NUMBER: AU PQ6457  
;; PRIOR FILING DATE: 2000-03-24  
;; NUMBER OF SEQ ID NOS: 128  
;; SOFTWARE: PatentIn version 3.0  
;; SEQ ID NO 126  
;; LENGTH: 337  
;; TYPE: PRT  
;; ORGANISM: Human  
US-09-814-777A-126

Query Match 13.7%; Score 79.5; DB 10; Length 337;  
Best Local Similarity 29.5%; Pred. No. 9.1;  
Matches 33; Conservative 7; Mismatches 39; Indels 33; Gaps 6;  
QY 28 AAEVPAQGOQGPGRGEEAPRGVMAVPLLR-MEGAPAG-----PGGRTAACFSC--- 77  
Db 184 AAFPPPPAAPRTARWRPSAPPTAHRYVAGPRLLRSGPGGAQDRAPRARSLACTTAPWA 243  
QY 78 -----TSRCLSRRPWRKRSWSAGSCPG-----MPHLSPDQGR 108  
Db 244 RPARTPARC-RRRPRPRWRAPS-PGARRRSVGRGPRHVRVPVQLQPSAR 293

Search completed: October 7, 2003, 13:58:54

Tue Oct 7 15:52:29 2003

us-09-807-512-2.rapb

Page 6

Job time : 119.658 secs

-----

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:19:12 ; Search time 6.1962 Seconds  
(without alignments)  
170.726 Million cell updates/sec

Title: US-09-807-512-11  
Perfect score: 50  
Sequence: 1 MLMQAQALAF 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues  
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_76.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query %	Score	Match	Length	DB	ID	Description
1	35	70.0	308	2	D83629		probable permease
2	34	68.0	207	2	H71671		ribosomal protein
3	34	68.0	310	2	D59098		hypothetical prote
4	34	68.0	388	2	AF1521		hypothetical prote
5	34	68.0	837	2	T48407		hypothetical prote
6	33	66.0	272	2	C91147		dehydroshikimate r
7	33	66.0	272	2	G85992		shikimate 5-dehydr
8	33	66.0	272	2	S00252		hypothetical prote
9	33	66.0	333	2	F84775		ubiquinol oxidase
10	33	66.0	665	2	H87468		C4-dicarboxylate t
11	32	64.0	96	1	I64010		protein gp21 [Bact
12	32	64.0	105	2	AB1360		PTS system, mannit
13	32	64.0	145	2	D84131		conserved hypothet
14	32	64.0	205	2	G82358		lysine decarboxyla
15	32	64.0	215	2	T48554		hypothetical prote
16	32	64.0	292	2	F64486		DNA-damage-inducib
17	32	64.0	363	2	E86878		chromosome segrega
18	32	64.0	1225	2	A49464		hypothetical prote
19	32	64.0	2697	2	T25444		2,4-dihydroxyhept-
20	31	62.0	263	2	AE0631		delta-aminolevulin
21	31	62.0	326	2	AD2353		3-Ketoacyl-CoA thi
22	31	62.0	398	2	F69266		hypothetical prote
23	31	62.0	425	2	T01228		minor capsid prote
24	31	62.0	439	1	VHBPCL		hypothetical prote
25	31	62.0	529	2	T23740		cytochrome-c oxida
26	31	62.0	554	2	S03809		virulence factor-r
27	31	62.0	555	2	E75513		signal peptidase I
28	31	62.0	636	2	G86504		signal peptidase I
29	31	62.0	636	2	F72118		

30	31	62.0	899	2	A84421	probable ATP-depen
31	31	62.0	922	2	S37871	hypothetical prote
32	31	62.0	4131	2	T21085	hypothetical prote
33	30	60.0	48	2	T22176	hypothetical prote
34	30	60.0	93	2	C75396	hypothetical prote
35	30	60.0	118	2	T16216	hypothetical prote
36	30	60.0	133	2	S69520	hypothetical prote
37	30	60.0	154	2	H64562	hypothetical prote
38	30	60.0	238	2	AH0918	conserved hypothet
39	30	60.0	274	2	D95165	NH(3)-dependent NA
40	30	60.0	274	2	C98031	NAD synthase (glut
41	30	60.0	275	1	PNEBT	beta-lactamase (EC
42	30	60.0	307	2	F95859	probable oligopept
43	30	60.0	315	2	AC3168	hypothetical prote
44	30	60.0	374	2	AF3071	hypothetical prote
45	30	60.0	374	2	C98215	oligopeptide ABC t

ALIGNMENTS

RESULT 1

D83629  
probable permease of ABC transporter PA0138 [imported] - Pseudomonas aeruginosa (stra  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: D83629  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa  
A:Reference number: A82950; MUID:20437337; PMID:10984043  
A:Accession: D83629  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-308 <STO>  
A:Cross-references: GB:AE004451; GB:AE004091; NID:g9945958; PIDN:AAG03528.1; GSPDB:GN  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA0138  
C:Superfamily: probable ribose ABC transporter rbsC-2

Query Match 70.0%; Score 35; DB 2; Length 308;

Best Local Similarity 63.6%; Pred. No. 13;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MLMQAQALAF 11

|||:|||:|

Db 137 MLFAQDALVYL 147

RESULT 2

H71671  
ribosomal protein L4 - Rickettsia prowazekii  
C:Species: Rickettsia prowazekii  
C:Date: 21-Nov-1998 #sequence\_revision 21-Nov-1998 #text\_change 03-Nov-2000  
C:Accession: H71671  
R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichteritz-Ponten, T.; Alsmark  
Nature 396, 133-140, 1998  
A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.  
A:Reference number: A71630; MUID:99039499; PMID:9823893  
A:Accession: H71671  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-207 <AND>  
A:Cross-references: GB:AJ235272; GB:AJ235269; NID:g3861033; PIDN:CAA15098.1; PID:g386  
A:Experimental source: strain Madrid E  
C:Genetics:  
A:Gene: rplD; RP658  
C:Superfamily: Escherichia coli ribosomal protein L4

Query Match 68.0%; Score 34; DB 2; Length 207;

Tue Oct 7 15:52:19 2003

Best Local Similarity 54.5%; Pred. No. 14;  
Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLMAQEALAF 11  
:||||:|  
Db 192 VLLSQEAVSFL 202

## RESULT 3

D59098 hypothetical protein pX01-60 - Bacillus anthracis virulence plasmid pX01

C:Species: Bacillus anthracis  
C:Date: 12-Nov-1999 #sequence\_revision 12-Nov-1999 #text\_change 16-Feb-2001  
C:Accession: D59098  
R:Okimaka, R.T.; Cloud, K.; Hampton, O.; Hoffmaster, A.R.; Hill, K.K.; Keim, P.; Koehler, J. Bacteriol. 181, 6509-6515, 1999  
A:Title: Sequence and organization of pX01, the large Bacillus anthracis plasmid harbored by B. anthracis strain Ames.  
A:Reference number: A59091; MUID:99445483; PMID:10515943  
A:Accession: D59098  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-310 <OKI>  
A:Cross-references: GB:AF065404; NID:g4894216; PIDN:AAD32364.1; PID:g4894276  
A:Experimental source: strain Sterne  
C:Genetics:  
A:Gene: pX01-60  
A:Genome: plasmid  
C:Superfamily: hypothetical protein pX01-60

Query Match 68.0%; Score 34; DB 2; Length 310;  
Best Local Similarity 63.6%; Pred. No. 21;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MLMAQEALAF 11  
:||||:|  
Db 263 MLMSKEAYGFL 273

## RESULT 4

AF1521 hypothetical protein lin0710 [Imported] - Listeria innocua (strain Clip11262)

C:Species: Listeria innocua  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AF1521  
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, H.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.; Science 294, 849-852, 2001  
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; May, O.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AF1521  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-388 <GLA>  
A:Cross-references: GB:AL592022; PIDN:CAC95942.1; PID:g16413162; GSPDB:GN00178  
A:Experimental source: strain Clip11262  
C:Genetics:  
A:Gene: lin0710

Query Match 68.0%; Score 34; DB 2; Length 388;  
Best Local Similarity 54.5%; Pred. No. 27;  
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLMAQEALAF 11  
:||||:|  
Db 346 MLLAQEMGFL 356

## RESULT 5

T48407 hypothetical protein F8F6.10 - Arabidopsis thaliana (fragment)

C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Jun-2000  
C:Accession: T48407  
R:Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.; submitted to the Protein Sequence Database, March 2000  
A:Reference number: Z24488  
A:Accession: T48407  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-837 <BEV>  
A:Cross-references: EMBL:AL162873  
A:Experimental source: cultivar Columbia; BAC clone F8F6  
C:Genetics:  
A:Map position: 5  
A:Note: F8F6.10  
C:Superfamily: Arabidopsis thaliana hypothetical protein T8H10.30

Query Match 68.0%; Score 34; DB 2; Length 837;  
Best Local Similarity 60.0%; Pred. No. 59;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLMAQEALAF 10  
:||||:|  
Db 345 MMAQDAVTF 354

## RESULT 6

C91147 dehydroshikimate reductase [Imported] - Escherichia coli (strain O157:H7, substrain R

C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001  
C:Accession: C91147  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.; Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.; DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genome analysis.  
A:Reference number: A99629; MUID:21156231; PMID:11256796  
A:Accession: C91147  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-272 <HAY>  
A:Cross-references: GB:BA000007; PIDN:BA037570.1; PID:g13363620; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain RMD 0509952  
C:Genetics:  
A:Gene: Ecs4147  
C:Superfamily: shikimate dehydrogenase; shikimate dehydrogenase homology

Query Match 66.0%; Score 33; DB 2; Length 272;  
Best Local Similarity 72.7%; Pred. No. 30;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MLMAQEALAF 11  
:||||:|  
Db 240 MLVAAHAFL 250

## RESULT 7

G85992 dehydroshikimate reductase [Imported] - Escherichia coli (strain O157:H7, substrain E

C:Species: Escherichia coli  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C:Accession: G85992  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May, J.; Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoudis, K.; Apoda, Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: G85992  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-272 <STO>  
A:Cross-references: GB:AE005174; NID:g12517908; PIDN:AAG58403.1; GSPDB:GN00145; UWGP  
A:Experimental source: strain O157:H7, substrain EDL933

## C:Genetics:

A:Gene: aroE  
A:Superfamily: shikimate dehydrogenase; shikimate dehydrogenase homology

Query Match 66.0%; Score 33; DB 2; Length 272;

Best Local Similarity 72.7%; Pred. No. 30;

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MLMQAQALAF 11

DB 240 MLVQAQAAHAF 250

## RESULT 8

S00252

shikimate 5-dehydrogenase (EC 1.1.1.25) - Escherichia coli (strain K-12)

C:Species: Escherichia coli

C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 01-Mar-2002

C:Accession: S00252; D65120

R:Anton, I.A.; Coggins, J.R.

Biochem. J. 249, 319-326, 1988

A:Title: Sequencing and overexpression of the Escherichia coli aroE gene encoding shikim

A:Reference number: S00252; MUID:88134021; PMID:3277621

A:Accession: S00252

A:Molecule type: DNA

A:Residues: 1-272 <EMBL>

A:Cross-references: EMBL:Y00710; NID:g40977; PIDN:CAA68700.1; PID:g40978

A:Note: Part of this sequence, including the amino end of the mature protein, was confir

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co

.A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: D65120

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-272 <BLAT>

A:Cross-references: GB:AE000406; GB:U00096; NID:g2367209; PIDN:AAC76306.1; PID:g1789675;

A:Experimental source: strain K-12, substrain MG1655

## C:Genetics:

A:Gene: aroE

A:Map position: 72 min

C:Superfamily: shikimate dehydrogenase; shikimate dehydrogenase homology

C:Keywords: oxidoreductase

F:1-272/Product: shikimate dehydrogenase #status experimental <MAT>

F:57-253/Domain: shikimate dehydrogenase homology <SKD>

Query Match 66.0%; Score 33; DB 2; Length 272;

Best Local Similarity 72.7%; Pred. No. 30;

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MLMQAQALAF 11

DB 240 MLVQAQAAHAF 250

## RESULT 9

F84775

hypothetical protein At2g36000 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C:Accession: F84775

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: F84775

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-333 <STO>

A:Cross-references: GB:AE002093; NID:g4510369; PIDN:AAD21457.1; GSPDB:GN00139

## C:Genetics:

A:Gene: At2g36000

A:Map position: 2

Query Match 66.0%; Score 33; DB 2; Length 333;

Best Local Similarity 77.8%; Pred. No. 37;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 MAQEAALAF 11

DB 1 MQEALSEL 9

## RESULT 10

H87468

ubiquinol oxidase (EC 1.10.3.-) subunit I [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 24-Aug-2001

C:Accession: H87468

R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: H87468

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-665 <STO>

A:Cross-references: GB:AE005673; NID:gl3423198; PIDN:AAK23748.1; GSPDB:GN00148

## C:Genetics:

A:Gene: CC1772

C:Superfamily: cytochrome-c oxidase chain I; cytochrome-c oxidase chain I homology

C:Keywords: copper; electron transfer; heme; iron; magnesium; membrane-associated com

F:114,429/Binding site: heme a iron (His) (axial ligands) #status predicted

F:292-296/Cross-link: 1'-histidyl-3'-tyrosine (His-Tyr) #status predicted

F:296/Binding site: oxygen (Tyr) #status predicted

Query Match 66.0%; Score 33; DB 2; Length 665;

Best Local Similarity 60.0%; Pred. No. 76;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLMQAQALAF 10

DB 86 MWRAQQAFAF 95

## RESULT 11

I64010

C4-dicarboxylate transport protein homolog HI0636 - Haemophilus influenzae (strain Rd

C:Species: Haemophilus influenzae

C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 28-Jul-2000

C:Accession: I64010

R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage

; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman

; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.

Science 269, 496-512, 1995

A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Vente

A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.

A:Reference number: A64000; MUID:95350630; PMID:7542800

A:Accession: I64010

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-96 <IGR>

A:Cross-references: GB:U32746; GB:L42023; NID:gl573626; PIDN:AAC22295.1; PID:gl573633

C:Superfamily: C4-dicarboxylate carrier protein

Query Match 64.0%; Score 32; DB 1; Length 96;

Best Local Similarity 100.0%; Pred. No. 17;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AQEALAF 10

DB 11111111

Tue Oct 7 15:52:19 2003

us-09-807-512-11.rpr

```

Db      42 AQEALAF 48

RESULT 12
AB1360
protein gp21 [Bacteriophage A118] [imported] - Listeria monocytogenes (strain EGD-e)
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AB1360
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluster, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlant,
A:Title: Comparative genomics of Listeria species
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AB1360
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-105 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CAD00360.1; PID:gl6411752; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo2282

Query Match      64.0%; Score 32; DB 2; Length 105;
Best Local Similarity 63.6%; Pred. No. 18;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 MLMQAEALAF 11
      :|: |||||
Db      89 LLLTQEAAL 99

RESULT 13
D84131
PTS system, mannitol-specific enzyme II, A component BH3852 [imported] - Bacillus halodurans
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: D84131
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: D84131
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-145 <STO>
A:Cross-references: GB:AP001520; GB:BA000004; NID:gi10176401; PIDN:BA07571.1; GSPDB:GN00
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH3852
C:Superfamily: phosphotransferase system mannitol-specific enzyme II factor III; phospho

Query Match      64.0%; Score 32; DB 2; Length 145;
Best Local Similarity 87.5%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      4 AQEALAF 11
      ||| ||||
Db     133 AQEVLAF 140

RESULT 14
G82358
conserved hypothetical protein VC0136 [imported] - Vibrio cholerae (strain N16961 serogr
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: G82358
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, R
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: G82358
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-205 <HEI>
A:Cross-references: GB:AF004104; GB:AE003852; NID:g9654534; PIDN:AAF93313.1; GSPDB:GN
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC0136
A:Map position: 1
C:Superfamily: hypothetical protein bl798

Query Match      64.0%; Score 32; DB 2; Length 205;
Best Local Similarity 77.8%; Pred. No. 37;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2 LMAQEALAF 10
      |::| ||||
Db      64 LVAQSALAF 72

RESULT 15
T48554
lysine decarboxylase-like protein - Arabidopsis thaliana
N:Alternate names: protein F14F18.120
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 17-Nov-2000
C:Accession: T48554
R:Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancro
submitted to the Protein Sequence Database, April 2000
A:Reference number: Z24490
A:Accession: T48554
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-215 <BEV>
A:Cross-references: EMBL:AL163812
A:Experimental source: cultivar Columbia; BAC clone F14F18
C:Genetics:
A:Map position: 5
A:Introns: 37/3; 70/2; 80/3; 114/1; 136/3; 184/3
A:Note: F14F18.120
C:Superfamily: yeast conserved hypothetical protein YJL055w

Query Match      64.0%; Score 32; DB 2; Length 215;
Best Local Similarity 77.8%; Pred. No. 39;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      3 MAQEALAF 11
      ||||| ||:
Db     102 MAQEAEAF 110

Search completed: October 7, 2003, 13:32:35
Job time : 9.1962 secs

```



Tue Oct 7 15:52:19 2003

```

DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Shikimate 5-dehydrogenase (EC 1.1.1.25).
GN AROE OR B3281.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RN SEQUENCE FROM N.A., AND SEQUENCE OF 1-30.
RP MEDLINE=88134021; PubMed=3277621;
RX Antton I.A., Coggins J.R.;
RT "Sequencing and overexpression of the Escherichia coli aroE gene
RL encoding shikimate dehydrogenase.";
RN Blochem. J. 249:319-326(1988).
RP [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RN X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX Maclean J., Campbell S.A., Pollock K., Chackrewarthi S., Coggins J.R.,
RA Laphorn A.J.;
RT "Crystallization and preliminary X-ray analysis of shikimate
RL dehydrogenase from Escherichia coli.";
RX Acta Crystallogr. D 56:512-515(2000).
CC -!- CATALYTIC ACTIVITY: Shikimate + NADP(+) -> 5-dehydroshikimate +
CC NADPH.
CC -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
CC fourth step.
CC -!- SUBUNIT: Monomer.
CC -!- SIMILARITY: Belongs to the shikimate dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Y00710; CAA68700.1; -
CC EMBL: U18997; AAA58078.1; -
CC EMBL: A8000406; AAC76306.1; -
CC PIR: S00252; S00252.
CC PDB: 1NYT; 1L-MAR-03.
CC HAMAP: MF_00222; aroE.
CC ECoGene; EG10077; aroE.
CC InterPro; IPR006152; Shikimate.
CC InterPro; IPR006151; Shikimate_DH.
CC Pfam; PF01488; Shikimate_DH; 1.
CC TIGRfams; TIGR00507; aroE; 1.
CC Aromatic amino acid biosynthesis; Oxidoreductase; NADP;
CC Complete proteome; 3D-structure;
CC SEQUENCE 272 AA; 29413 MW; 62D3797ECA1EC1E8 CRC64;
CC
CC Query Match 66.0%; Score 33; DB 1; Length 272;
CC Best Local Similarity 72.7%; Pred. No. 11;
CC Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
CC
CC QY 1 MLMAQEAALAF 11
CC Db 240 MLVQAQAHAFL 250
CC
CC RESULT 4

```



```

Y636_HAEIN
ID Y636_HAEIN STANDARD; PRT; 96 AA.
AC P44027;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein HI0636.
GN HI0636.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RD.";
RL Science 269:496-512(1995).
CC -!- SIMILARITY: BELONGS TO THE YFCZ/YIIS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U32746; AAC22295.1; -
DR PIR: I64010; I64010.
DR TIGR: HI0636; -
DR InterPro: IPR005272; Cons_hypoth743.
DR Pfam: PF04175; DUF406; 1.
DR TIGRFAMs: TIGR00743; TIGR00743; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 96 AA; 10966 MW; 7719C9CFD29587F CRC64;
-----
Query Match 64.0%; Score 32; DB 1; Length 96;
Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AQEALAF 10
DB 42 AQEALAF 48
|||||

RESULT 5
PTMA_BACHD STANDARD; PRT; 145 AA.
AC Q9K680;
DT 16-OCT-2001 (Rel. 40, Created)
DT 28-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE PTS system, mannitol-specific IIA component (ETIA-Mtl) (Mannitol-
DE permease IIA component) (Phosphotransferase enzyme II, A component)
DE (EC 2.7.1.69) (ETII-Mtl).
GN MTLF OR BH382.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C-125 / JCM 9153;

```

```

RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RT Horikoshi K.;
RT Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -!- FUNCTION: THIS IS A COMPONENT OF THE PHOSPHOENOLPYRUVATE-DEPENDENT
CC SUGAR PHOSPHOTRANSFERASE SYSTEM (PTS), A MAJOR CARBOHYDRATE ACTIVE
CC -TRANSPORT SYSTEM. THE IID DOMAINS CONTAIN THE SUGAR BINDING SITE
CC AND THE TRANSMEMBRANE CHANNEL. THE IIA DOMAIN CONTAINS THE PRIMARY
CC PHOSPHORYLATION SITE (THE DONOR IS PHOSPHO-HPR); IIA TRANSFERS ITS
CC PHOSPHORYL GROUP TO THE IIB DOMAIN WHICH FINALLY TRANSFERS IT TO
CC THE SUGAR (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: Protein N-phosphohistidine + sugar -> protein
CC histidine + sugar phosphate.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Contains 1 PTS_EIIA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP001520; BAB07571.1; -
DR PIR: D84131; D84131.
DR HSP; P00350; IAJA.
DR InterPro: IPR002178; PTS_EIIA_2.
DR Pfam: PF00359; PTS_EIIA_2; 1.
DR ProDom: PD001689; PTS_EIIA_2; 1.
DR PROSITE: PS00372; PTS_EIIA_2; 1.
KW Phosphotransferase system; Sugar transport; Transferase;
KW Phosphorylation; Complete proteome.
FT MOD_RES 64 64 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 145 AA; 16015 MW; 888B849FA9C9B7FD CRC64;
-----
Query Match 64.0%; Score 32; DB 1; Length 145;
Best Local Similarity 87.5%; Pred. No. 9.6;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 AQEALAF 11
DB 1133 AQEVLAF 140
|||||

RESULT 6
YE95_METJA STANDARD; PRT; 292 AA.
ID YE95_METJA
AC Q58890;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ1495.
GN MJ1495.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;

```

```
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RL Jannaschii." 1058-1073(1996).
CC -|- SIMILARITY: STRONG, TO M.JANNASCHII MJ0137.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U67590; AAB99506.1; -.
CC TIGR: MJ1495; -.
CC InterPro: IPR002825; DUF114.
CC Pfam: PF01972; DUF114; 1.
CC Hypothetical protein; Transmembrane; Complete proteome.
KW TRANSMEM 17 37 POTENTIAL.
FT SEQUENCE 292 AA; 33508 MW; B8A34C29B3A51246 CRC64;
SQ
Query Match 64.0%; Score 32; DB 1; Length 292;
Best Local Similarity 63.6%; Pred. No. 20;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 MLMQAQALAF 11
DB 62 MIHQEALTEL 72
I: |||||
RESULTS
ID DP04_LACIA STANDARD; PRT; 363 AA.
AC Q9CE21;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA polymerase IV (EC 2.7.7.7) (Pol IV).
DN DIB OR L12029.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Manger S., Jaillon O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403."
RL Genome Res. 11:731-753(2001).
CC -|- FUNCTION: Poorly processive, error-prone DNA polymerase involved
CC in untargeted mutagenesis. Copies undamaged DNA at stalled
CC replication forks, which arise in vivo from mismatched or
CC misaligned primer ends. These misaligned primers can be extended
CC by polIV. Exhibits no 3'-5' exonuclease (proofreading) activity.
CC May be involved in translesional synthesis, in conjunction with
CC the beta clamp from polIII (By similarity).
CC -|- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate
CC + {DNA}(N).
CC -|- COFACTOR: Binds two magnesium ions (By similarity).
CC -|- SUBUNIT: Monomer (By similarity).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -|- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-Y FAMILY.
CC -|- SIMILARITY: Contains 1 umuC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE006432; AAK06127.1; -.
CC PIR: E86878; E86878.
CC HAMAP: MF_01113; -. 1.
CC InterPro: IPR001126; UMUC_like.
CC PROSITE: PS00817; IMS; 1.
CC Transferrase; DNA-directed DNA polymerase; DNA replication; DNA repair;
KW DNA-binding; Mutator protein; Magnesium; Complete proteome.
FT DOMAIN 14 197 UMUC
FT SITE 23 23 SUBSTRATE DISCRIMINATION (BY SIMILARITY).
FT ACT_SITE 116 116 BY SIMILARITY.
FT METAL 18 18 MAGNESIUM (BY SIMILARITY).
FT METAL 115 115 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 363 AA; 40438 MW; 4E4DFA9C0CEC168B CRC64;
Query Match 64.0%; Score 32; DB 1; Length 363;
Best Local Similarity 54.5%; Pred. No. 25;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 1 MLMQAQALAF 11
DB 175 LIMPEEALEFL 185
I: |||||
RESULTS
ID SMCL_YEAST STANDARD; PRT; 1225 AA.
AC P32908;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Structural maintenance of chromosome 1 (DA-box protein SMCL1).
GN SMCL1 OR CHE10 OR YFL008W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A. AND MUTANTS SMCL1-1 AND SMCL1-2.
RX MEDLINE=94103320; PubMed=8276886;
RA Strunnikov A.V., Larionov V.L., Koshland D.;
RT "SMCL1: an essential yeast gene encoding a putative head-rod-tail
RT protein is required for nuclear division and defines a new ubiquitous
RT protein family."
RL J. Cell Biol. 123:1635-1648(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=95400292; PubMed=7670463;
RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,
RA Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA Yamazaki M., Tashiro H., Eki T.;
RT "Analysis of the nucleotide sequence of chromosome VI from
RT Saccharomyces cerevisiae."
RL Nat. Genet. 10:261-268(1995).
RN [3]
RP IDENTIFICATION IN A COHESIN COMPLEX WITH SMC3; IRR1 AND MCD1, AND
RP INTERACTION OF THE COHESIN COMPLEX WITH SCC2.
RX PubMed=9990856;
RA Toth A., Ciosk R., Uhlmann F., Galova M., Schleiffer A., Nasmyth K.;
RT "Yeast cohesin complex requires a conserved protein, Eco1p(Ctf7), to
RT establish cohesion between sister chromatids during DNA replication."
RL Genes Dev. 13:320-333(1999).
RN [4]
RP IDENTIFICATION IN A COHESIN COMPLEX WITH SMC3; MCD1 AND IRR1, AND
RP STRUCTURE.
RX PubMed=11983169;
RA Haering C.H., Loewe J., Hochwagen A., Nasmyth K.;
RT "Molecular architecture of SMC proteins and the yeast cohesin
RT complex."
RL Mol. Cell 9:773-788(2002).
```



```

or send an email to license@isb-sib.ch).
-----
CC EMBL; X05829; CAA29274.1; -.
CC PIR; S03809; S03809.
DR HSSP; P98002; IAR1.
DR InterPro; IPRO00883; COX1.
DR Pfam; PF00115; COX1; 1.
DR PRINTS; PRO1165; CYC oxidase I.
DR PROSITE; PS00077; COXI; 1.
KW Oxidoreductase; Heme; Copper; Transmembrane; Inner membrane;
KW Respiratory chain; Hydrogen ion transport.
FT TRANSFAM 26 56 BY SIMILARITY.
FT TRANSMEM 81 118 BY SIMILARITY.
FT TRANSMEM 127 148 BY SIMILARITY.
FT TRANSMEM 175 203 BY SIMILARITY.
FT TRANSMEM 215 248 BY SIMILARITY.
FT TRANSMEM 260 295 BY SIMILARITY.
FT TRANSMEM 301 319 BY SIMILARITY.
FT TRANSMEM 331 359 BY SIMILARITY.
FT TRANSMEM 367 390 BY SIMILARITY.
FT TRANSMEM 399 425 BY SIMILARITY.
FT TRANSMEM 436 463 BY SIMILARITY.
FT TRANSMEM 478 508 BY SIMILARITY.
FT METAL 91 IRON (HEME A AXIAL LIGAND) (BY SIMILARITY).
FT METAL 273 COPPER B (BY SIMILARITY).
FT METAL 277 COPPER B (BY SIMILARITY).
FT METAL 322 COPPER B (BY SIMILARITY).
FT METAL 323 COPPER B (BY SIMILARITY).
FT METAL 406 IRON (HEME A3 AXIAL LIGAND) (BY SIMILARITY).
FT METAL 408 IRON (HEME A AXIAL LIGAND) (BY SIMILARITY).
FT DISULFID 63 77 BY SIMILARITY.
FT CROSSLINK 273 277 1'-histidyl-3'-tyrosine (By similarity).
SQ SEQUENCE 554 AA; 62012 MW; 58AD591FBDCFE794 CRC64;
Query Match 62.0%; Score 31; DB 1; Length 554;
Best Local Similarity 54.5%; Pred. No. 63;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0
Oy 1 MLMQAQLAFL 11
Db 305 MYLMAAIAFL 315

RESULT 12
YKF0_YEAST STANDARD; PRT; 922 AA.
ID YKF0_YEAST STANDARD; PRT; 922 AA.
AC AC
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DE 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 103.1 kDa protein in NUP120-CSE4 intergenic region.
GN YKL050C OR YKL301 OR YKL263.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
NCBI_TaxID=4932;
[1]
RN RP SEQUENCE OF 1-452 FROM N.A.
RC STRAIN=S288c;
RX MEDLINE=94205268; PubMed=8154189;
RA Purnelle B., Skala J., van Dyck L., Tettelin H., Goffeau A.;
RT "The sequence of a 17.5 kb DNA fragment on the left arm of yeast chromosome XI identifies the protein kinase gene ELM1, the DNA primase gene PR12, a new gene encoding a putative histone and seven new open reading frames.";
RL Yeast 9:1379-1384(1993).
[2]
RN RP SEQUENCE OF 410-922 FROM N.A.
RC STRAIN=S288c;
RX MEDLINE=943378723; PubMed=8091862;
```

```

RA Rasmussen S.W.;
RT "Sequence of a 28.6 kb region of yeast chromosome XI includes the
RT FBAl and tOA2 genes, an open reading frame (ORF) similar to a
RT translationally controlled tumour protein, one ORF containing motifs
RT also found in plant storage proteins and 13 ORFs with weak or no
RT homology to known proteins.";
RL Yeast 10:S63-S68(1994).
CC -!- SIMILARITY: TO YEAST YMR031C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X75781; CAA53420.1; -
DR EMBL; X71621; -; NOT_ANNOTATED_CDS.
DR EMBL; Z28050; CAA81885.1; -
DR PIR; S37871; S37871.
DR SGD; S0001533; YKL050C.
KW Hypothetical protein.
SQ SEQUENCE 922 AA; 103143 MW; 362EE9A3F642DC8B CRC64;

Query Match 62.0%; Score 31; DB 1; Length 922;
Best Local Similarity 77.8%; Pred. No. 1.1e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LMAQALAF 10
DB 173 LAAQALAF 181

RESULT 13
Y014_BPH1 STANDARD; PRT; 133 AA.
AC P51716;
DT 01-OCT-1996 (Rel. 34; Created)
DT 01-OCT-1996 (Rel. 34; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Hypothetical 14.9 kDa protein in REP-HOL intergenic region (ORF14).
OS Bacteriophage HPI.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
OX NCBI_TaxID=10690;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HP1C1;
RX MEDLINE=96279738; PubMed=8710508;
RA Esposito D., Fitzmaurice W.P., Benjamin R.C., Goodman S.D.,
RA Waldman A.S., Scocca J.J.
RT "The complete nucleotide sequence of bacteriophage HPI DNA.";
RL Nucleic Acids Res. 24:2360-2368(1996).
CC -!- SIMILARITY: TO E.COLI YDCQ.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U24159; AAB09199.1; -
DR PIR; S69520; S69520.
DR InterPro; IPR005357; UPF0150.
DR Pfam; PF03681; UPF0150; 1.
KW Hypothetical protein.
SQ SEQUENCE 133 AA; 14922 MW; 8B0A669D7F018922 CRC64;

Query Match 60.0%; Score 30; DB 1; Length 133;
Best Local Similarity 60.0%; Pred. No. 24;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 MLMAQALAF 10
DB 35 MLNAKEAISF 44

RESULT 14
NADE_STRP3 STANDARD; PRT; 274 AA.
AC Q8KG04;
DT 28-FEB-2003 (Rel. 41; Created)
DT 28-FEB-2003 (Rel. 41; Last sequence update)
DT 15-SEP-2003 (Rel. 42; Last annotation update)
DE NH(3)-dependent NAD(+) synthetase (EC 6.3.5.1).
GN NADE OR SPW3_1392 OR SP0470.
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=198466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS315 / Serotype M3;
RX MEDLINE=22133808; PubMed=12122206;
RA Beres S.B., Sylva G.L., Barbian K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.
RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
RT phage-encoded toxins, the high-virulence phenotype, and clone
RT emergence.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SSI-1 / Serotype M3;
RA Nakagawa I., Kurokawa K., Nakata M., Tomiyasu Y., Yamashita A.,
RA Yamazaki K., Okahashi N., Kawabata S., Yasunaga T., Hattori M.,
RA Hayashi H., Hamada S.;
RT "The genome of invasive Streptococcus pyogenes; a comparative analysis
RT of S. pyogenes SST-1, SF370 and MGAS8232.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ATP + deamido-NAD(+) + NH(3) + H(2)O -> AMP +
CC diphosphate + NAD(+).
CC -!- PATHWAY: NAD biosynthesis.
CC -!- SIMILARITY: BELONGS TO THE NAD SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE014163; AAM7999.1; -
DR EMBL; AP005142; BAC63565.1; -
DR HAMAP; MF_00193; -; 1.
DR InterPro; IPR003694; NAD_synthase.
DR Pfam; PF02540; NAD_synthase; 1.
DR TIGRFAMs; TIGR00552; nade; 1.
KW Ligase; NAD; ATP-binding; Complete proteome.
KW NP_BIND 46 53
FT ACT_SITE 48 48 BY SIMILARITY.
SQ SEQUENCE 274 AA; 30180 MW; C52B01B7F9D40B87 CRC64;

Query Match 60.0%; Score 30; DB 1; Length 274;
Best Local Similarity 75.0%; Pred. No. 50;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 AQEALAF 11
DB 94 AOKALAFI 101

```

```
RESULT 15
NADE_STRP8 STANDARD; PRT; 274 AA.
AC Q8NZY8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE NH(3)-dependent NAD(+) synthetase (EC 6.3.5.1).
GN NADE OR SPYM18_1663.
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=186103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MGAS8232 / Serotype M18;
RX MEDLINE-21927593; PubMed-11917108;
RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F., Zhang Q.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks";
RT Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
CC -|- CATALYTIC ACTIVITY: ATP + deamido-NAD(+) + NH(3) + H(2)O = AMP +
CC diposphate + NAD(+).
CC -|- PATHWAY: NAD biosynthesis.
CC -|- SIMILARITY: BELONGS TO THE NAD SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE010078; AAL98207.1; -.
DR HAMAP; MF_00193; -.
DR InterPro; IPR003694; NAD_synthase.
DR Pfam; PF02540; NAD_synthase; 1.
DR TIGRFAMs; TIGR00552; nade; 1.
KW Ligase; NAD; ATP-binding; Complete proteome.
FT NP_BIND 46 53 ATP (BY SIMILARITY).
FT ACT_SITE 48 48 BY SIMILARITY.
SQ SEQUENCE 274 AA; 30208 MW; DA7CD7F2004BB3C7 CRC64;

Query Match 60.0%; Score 30; DB 1; Length 274;
Best Local Similarity 75.0%; Pred. No. 50;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 AQENLAFL 11
DB 94 AQKALAFI 101

Search completed: October 7, 2003, 13:26:23
Job time : 5.20253 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:04:57 ; Search time 16.5823 Seconds  
(without alignments)  
95.721 Million cell updates/sec

Title: US-09-807-512-12  
Perfect score: 45  
Sequence: 1 LMAQELAF 10  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues  
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*  
1: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.\*  
2: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.\*  
3: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.\*  
4: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.\*  
5: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.\*  
6: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.\*  
7: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.\*  
8: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.\*  
9: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.\*  
10: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.\*  
11: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.\*  
12: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.\*  
13: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.\*  
14: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.\*  
15: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.\*  
16: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.\*  
17: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.\*  
18: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.\*  
19: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.\*  
20: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.\*  
21: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.\*  
22: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.\*  
23: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.\*  
24: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	100.0	10	21	CTL epitope-2 of h
2	45	100.0	11	21	CTL epitope-1 of h
3	45	100.0	11	23	Tumour antigen LAG
4	45	100.0	30	23	Human NYN501b segm
5	45	100.0	58	20	Human cancer antig
6	45	100.0	58	21	Human tumour antig
7	45	100.0	58	23	Human NYN501b cons
8	45	100.0	109	21	Human CTL-recognis
9	45	100.0	3541	23	Human melanoma spe

10	41	91.1	10	23	ABG79072	Human CAMEL class
11	34	75.6	222	23	ABP73965	Candida albicans e
12	34	75.6	577	23	ABU65209	Human NOV121a prot
13	33	73.3	320	21	AAG22016	Arabidopsis thalia
14	33	73.3	761	22	AAB47134	CDIFF-15, Incyte I
15	33	73.3	891	23	AAE22231	Human schlafen C p
16	32	71.1	215	23	ABB93441	Herbicidally activ
17	32	71.1	552	22	AAAB5484	Pfam consensus seq
18	31	68.9	143	23	ABB93600	Herbicidally activ
19	31	68.9	395	23	ABJ11257	Yeast selected int
20	31	68.9	395	23	ABG77355	Selected Interacti
21	31	68.9	398	22	AAB48047	Signal transductio
22	31	68.9	398	22	AAB48048	Signal transductio
23	31	68.9	636	20	AAAY34710	Amino acid sequenc
24	30	66.7	63	23	ABP34124	Human ORF3097 prot
25	30	66.7	92	20	AAAY13124	Human secreted pro
26	30	66.7	105	23	ABR49797	Listeria monocytog
27	30	66.7	234	22	AAU13426	Human diagnostic a
28	30	66.7	257	24	ABR47519	Breast cancer asso
29	30	66.7	274	21	AAAB21123	Streptococcus pneu
30	30	66.7	274	22	AAU37693	Streptococcus pneu
31	30	66.7	274	24	ABU01867	S. pneumoniae type
32	30	66.7	288	23	ABP26883	Streptococcus poly
33	30	66.7	363	23	ABB55383	Lactococcus lactis
34	30	66.7	463	21	AAAY50930	Human fetal brain
35	30	66.7	463	21	AAAY44276	Human nucleic acid
36	30	66.7	590	22	AAAM78688	Human protein SEQ
37	30	66.7	873	23	ABP30006	Streptococcus poly
38	29	64.4	69	22	AAO00664	Human polypeptide
39	29	64.4	85	22	ABG51931	Human liver peptid
40	29	64.4	85	22	ABB31862	Peptide #4513 enco
41	29	64.4	85	22	ABB37098	Peptide #4604 enco
42	29	64.4	85	22	ABB22407	Protein #4406 enco
43	29	64.4	85	22	AAAM57820	Human brain expres
44	29	64.4	85	22	AAAM70234	Human bone marrow
45	29	64.4	85	22	AAAM18060	Peptide #4494 enco

ALIGNMENTS

RESULT 1

AAAY70856  
ID AAAY70856 standard; peptide; 10 AA.

XX AAAY70856;

XX 31-JUL-2000 (first entry)

DT CTL epitope-2 of human CAMEL protein.

DE CAMEL; CTL-recognised Antigen on MELANOMA; cytotoxic T lymphocyte; CTL;  
tumour-associated antigen; LAGE-1; NY-ESO-1; anticancer; melanoma;  
human; cancer; immunotherapy; immunogenic peptide; immune response.

XX Homo sapiens.

XX WO2000023584-A1.

XX 27-APR-2000.

XX 15-OCT-1999; 99WO-EP07832.

XX 16-OCT-1998; 98EP-0119583.

XX (BOEH ) BOEHRINGER INGELHEIM INT GMBH.  
XX (UYHO-) UNIV HOSPITAL LEIDEN.

XX Schrier PI, Aarnoudse CA, Heider K, Klade C;

XX WPI; 2000-339685/29.

XX Tumor-associated antigen useful for cancer immunotherapy is encoded by

CC different open reading frame (ORF-1). It shows strong homology with  
CC NY-ESO-1, a melanoma specific tumour antigen. The tumour-associated  
CC antigen displayed on melanoma cells is recognised by cytotoxic T  
CC lymphocytes. CAMEL is expressed in tumour cell lines, tumour tissues  
CC (e.g. breast and lung) and in restricted number of healthy tissues. This  
CC sequence has anticancer activity. CAMEL tumour antigen and immunogenic  
CC peptides derived from it are useful for cancer immunotherapy. They have  
CC the potential to induce an immune response, by eliciting a CTL response.  
CC The DNA molecule is used to construct recombinant or fusion proteins.  
XX  
SQ Sequence 11 AA;  
Query Match 100.0%; Score 45; DB 21; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.024;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LMAQEALAFLL 10  
| | | | | | | | | |  
DB 2 LMAQEALAFLL 11  
RESULT 3  
ABG66804  
ID ABG66804 standard; Peptide; 11 AA.  
XX  
AC ABG66804;  
XX  
DT 24-SEP-2002 (first entry)  
XX  
DE Tumour antigen LAGE/CAMEL, HLA-A2 epitope.  
XX  
KW Beta-2 microglobulin; beta-2m; cytotoxic T lymphocyte; CTL; HLA;  
KW human leukocyte antigen; fusion protein; epitope; cytostatic; tumour;  
KW gastrointestinal tumour; colorectal cancer; gastro-oesophageal cancer;  
KW liver cancer; biliary tract cancer; pancreatic cancer; vaccine;  
KW prostatic cancer; testicular cancer; lung cancer; breast cancer;  
KW malignant melanoma; mesothelioma; brain tumour; ovarian cancer;  
KW uterine cancer; cervical cancer; head and neck cancer; bladder cancer;  
KW Kaposi's sarcoma; renal carcinoma; leukaemia; lymphoma;  
KW acquired immunodeficiency syndrome; AIDS-related lymphoma.  
XX  
OS Homo sapiens.  
XX  
PN WO200236146-A2.  
XX  
PD 10-MAY-2002.  
XX  
PF 01-NOV-2001; 2001WO-GB04844.  
XX  
PR 02-NOV-2000; 2000GB-0026812.  
XX  
PA (ISIS-) ISIS INNOVATION LTD.  
XX  
PI Tafuro S, Meier U, McMichael AJ, Bell JI, Layton G, Hunter M;  
XX  
DR WPI; 2002-508108/54.  
XX  
PT New polynucleotide capable of expressing an epitope-beta2m fusion  
PT protein useful for generating cytotoxic T lymphocyte responses against  
PT a tumour and in restoring antigen presentation in the tumour of a host  
PT  
XX  
PS Disclosure; Page 25; 46pp; English.  
XX  
CC The invention relates to a new polynucleotide capable of expressing an  
CC epitope-beta\_2m fusion protein useful for generating cytotoxic T  
CC lymphocyte (CTL) responses against a tumour or in restoring antigen  
CC presentation in the tumour of a host. Also included are a polynucleotide  
CC capable of expressing an epitope-beta\_2m fusion protein in combination  
CC with a vaccination agent that stimulates a CTL response against the  
CC epitope of the fusion protein for simultaneous, separate or sequential  
CC use in the treatment of cancer and a method of treating a tumour by  
CC administering a capable of expressing an epitope-beta\_2m fusion protein,

PT the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA -  
XX  
PS Claim 5; Page 34; 73pp; English.  
XX  
CC The present sequence is an immunogenic peptide of human tumour-associated  
CC antigen CAMEL (cytotoxic T lymphocytes (CTL)-recognised Antigen on  
CC MELANOMA). This peptide is a CTL epitope, that has the ability to elicit  
CC a CTL response. It corresponds to residues 2-11 of the CAMEL protein.  
CC CAMEL protein is encoded by the LAGE-1 gene, a tumour-specific antigen.  
CC It is different from the LAGE-1 protein, since it is translated from a  
CC different open reading frame (ORF-1). It shows strong homology with  
CC NY-ESO-1, a melanoma specific tumour antigen. The tumour-associated  
CC antigen displayed on melanoma cells is recognised by cytotoxic T  
CC lymphocytes. CAMEL is expressed in tumour cell lines, tumour tissues  
CC (e.g. breast and lung) and in restricted number of healthy tissues. This  
CC sequence has anticancer activity. CAMEL tumour antigen and immunogenic  
CC peptides derived from it are useful for cancer immunotherapy. They have  
CC the potential to induce an immune response, by eliciting a CTL response.  
CC The DNA molecule is used to construct recombinant or fusion proteins.  
XX  
SQ Sequence 10 AA;  
Query Match 100.0%; Score 45; DB 21; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.022;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LMAQEALAFLL 10  
| | | | | | | | | |  
DB 1 LMAQEALAFLL 10  
RESULT 2  
AAY70855  
ID AAY70855 standard; peptide; 11 AA.  
XX  
AC AAY70855;  
XX  
DT 31-JUL-2000 (first entry)  
XX  
DE CTL epitope-1 of human CAMEL protein.  
XX  
KW CAMEL; CTL-recognised Antigen on MELANOMA; cytotoxic T lymphocyte; CTL;  
KW tumour-associated antigen; LAGE-1; NY-ESO-1; anticancer; melanoma;  
KW human; cancer; immunotherapy; immunogenic peptide; immune response.  
XX  
OS Homo sapiens.  
XX  
PN WO200023584-A1.  
XX  
PD 27-APR-2000.  
XX  
PF 15-OCT-1999; 99WO-EP07832.  
XX  
PR 16-OCT-1998; 98EP-0119583.  
XX  
PA (BOEH ) BOEHRINGER INGELHEIM INT GMBH.  
XX  
PI (UYHO-) UNIV HOSPITAL LEIDEN.  
XX  
PI Schrier PI, Aarnoudse CA, Heider K, Klade C;  
XX  
DR WPI; 2000-339685/29.  
XX  
PT Tumor-associated antigen useful for cancer immunotherapy is encoded by  
PT the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA -  
XX  
PS Claim 4; Page 34; 73pp; English.  
XX  
CC The present sequence is an immunogenic peptide of human tumour-associated  
CC antigen CAMEL (cytotoxic T lymphocytes (CTL)-recognised Antigen on  
CC MELANOMA). This peptide is a CTL epitope, that has the ability to elicit  
CC a CTL response. It corresponds to residues 1-11 of the CAMEL protein.  
CC CAMEL protein is encoded by the LAGE-1 gene, a tumour-specific antigen.  
CC It is different from the LAGE-1 protein, since it is translated from a



CC and optionally a vaccination agent that stimulates a CTL response against  
 CC the epitope of the fusion protein. The polynucleotide is useful for  
 CC generating CTL responses against tumours, for restoring antigen  
 CC presentation in the tumour, and subsequently for treating cancers, such  
 CC as gastrointestinal tumour, prostatic, testicular, lung or breast cancer,  
 CC malignant melanoma, mesothelioma, brain tumour, ovarian cancer, uterine  
 CC cancer including cervical cancer, cancer of the head and neck, bladder  
 CC cancer, Kaposi's sarcoma, AIDS (acquired immunodeficiency syndrome)-  
 CC related Kaposi's sarcoma, sarcomas, osteosarcoma, renal carcinoma, and  
 CC haematopoietic malignant tumours such as leukaemia and lymphoma.  
 CC The epitope is an HLA (human leukocyte antigen) peptide derived from a  
 CC viral or tumour antigen. The present sequence is a tumour HLA epitope  
 CC used in the fusion proteins of the invention.

XX Sequence 11 AA;

Query Match 100.0%; Score 45; DB 23; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.024;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMAQEALAF 10  
 | | | | | | | |  
 Db 2 LMAQEALAF 11

#### RESULT 4

AAU85114  
 ID AAU85114 standard; Peptide; 30 AA.

AC AAU85114;

DT 08-MAY-2002 (first entry)

DE Human NYN501b segment 1.

KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;  
 KW viral infection; human immunodeficiency virus; melanoma;  
 KW bacterial infection; Salmonella; Legionella; parasitic infection;  
 KW Trypanosoma; Toxoplasma; Giardia.

XX Homo sapiens.

XX WO200190197-A1.

XX 29-NOV-2001.

XX 25-MAY-2001; 2001WO-AU00622.

XX 26-MAY-2000; 2000AU-0007761.

XX (AUSU) UNIV AUSTRALIAN NAT.

XX Thomson SA, Ramshaw IA;

XX WPI: 2002-147575/19.

XX N-PSDB; ABK36934.

XX New synthetic polypeptides having several different segments of at  
 PT least one parent polypeptide linked together differently compared to  
 PT the linkage in the parent polypeptide, for inducing immune response  
 PT against a pathogen or cancer.

XX Example 3; Fig 27; 364pp; English.

XX The invention relates to a new synthetic polypeptide (I) comprising  
 CC several different segments of at least one parent polypeptide linked  
 CC together in a different relationship relative to their linkage in the  
 CC parent polypeptide to impede, abrogate or otherwise alter at least one  
 CC function associated with the parent polypeptide and for inducing an  
 CC immune response against a pathogen or cancer. Also included are a  
 CC synthetic polynucleotide encoding and a computer system for  
 CC designing the synthetic polypeptides. The synthetic polypeptides and  
 CC polynucleotides are referred to as a Savine. The synthetic polypeptide is

CC useful for modulating immune responses preferably directed against a  
 CC pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix,  
 CC colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone  
 CC liver, oesophagus, brain, testicle, uterus), as potentiating agents.  
 CC Compositions comprising the polypeptide may be used in the treatment or  
 CC prophylaxis against viral (such as infections caused by HIV (human  
 CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis  
 CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial  
 CC (e.g., infections caused by Neisseria, Meningococcus, Haemophilus,  
 CC Salmonella, Streptococcus, Legionella and Mycobacterium or parasitic  
 CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,  
 CC Trypanosoma, Toxoplasma and Giardia) infections. The present  
 CC sequence is a peptide derived from a parent protein used to  
 CC construct a savine of the invention.

SQ Sequence 30 AA;

Query Match 100.0%; Score 45; DB 23; Length 30;  
 Best Local Similarity 100.0%; Pred. No. 0.066;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMAQEALAF 10  
 | | | | | | | |  
 Db 4 LMAQEALAF 13

#### RESULT 5

AAU05966

ID AAU05966 standard; Protein; 58 AA.

XX AAU05966;

XX 16-AUG-1999 (first entry)

XX Human cancer antigen NY ESO-1/CAG-3 ORF2 protein.

XX NY ESO-1/CAG-3 gene; CAG-3 gene; cancer antigen; human; leukaemia;  
 KW non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
 KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
 KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
 KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
 KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
 KW vaccine; ORF2.

XX Homo sapiens.

XX WO9918206-A2.

XX 15-APR-1999.

XX 21-SEP-1998; 98WO-US19609.

XX 08-OCT-1997; 97US-0061428.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Rosenberg SA, Wang RF;

XX WPI: 1999-277270/23.

XX N-PSDB; AAX58599.

XX Cancer antigen NY ESO1/CAG-3

XX Claim 5; Fig 3A; 88pp; English.

XX The present sequence represents the ORF2 protein encoded by  
 CC open reading frame 2 of the human ESO-1/CAG-3 (or CAG-3) gene.  
 CC CAG-3 is a new and potent tumour antigen capable of eliciting an  
 CC antigen specific immune response by T cells. Cancer peptides  
 CC comprising ORF2, ORF1 (see AAU05965), portions of these peptides and  
 CC their variants (see AAU05967-87), are useful as cancer vaccines that  
 CC protect the recipient from development of cancer. The invention  
 CC provides: vectors and host cells (also useful as vaccines); a

CC method of diagnosis of cancer or precancer; a transgenic animal;  
CC antisense oligonucleotides that inhibit expression of the cancer  
CC peptide or tumour antigen; antibodies reacting with the CAG-3  
CC cancer peptide, useful in diagnostic and detection assays; and  
CC methods for preventing or inhibiting cancer by administering a  
CC cancer peptide, with or without an HLA molecule. The cancer  
CC peptides form part of, or are derived from, cancers such as  
CC primary or metastatic melanoma, thymoma, lymphoma, sarcoma, lung  
CC cancer, liver cancer, leukaemia, uterine cancer, cervical cancer,  
CC bladder cancer, kidney cancer and adenocarcinomas such as breast,  
CC prostate, ovarian, pancreatic and thyroid cancers. Melanoma is  
CC treated by inducing cancer-specific T cells in vitro for subsequent  
CC return to a patient.

XX  
SQ Sequence 58 AA;  
Query Match 100.0%; Score 45; DB 20; Length 58;  
Best Local Similarity 100.0%; Pred. No. 0.13;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMAQALAF 10  
Db 2 LMAQALAF 11  
|||||

RESULT 6  
AAV70863  
ID AAV70863 standard; Protein: 58 AA.  
XX  
AC AAV70863;  
XX  
DT 31-JUL-2000 (first entry)  
XX  
DE Human tumour antigen, NY-ESO-1 short variant protein.  
XX  
KW NY-ESO-1; CAMEL; CTL-recognised Antigen on MELANOMA; human; cancer; CTL;  
KW cytotoxic T lymphocyte; tumour-associated antigen; LAGE-1; anticancer;  
KW melanoma; immunotherapy; immune response; variant.  
XX  
OS Homo sapiens.  
XX  
PN WO200023584-A1.  
XX  
PD 27-APR-2000.  
XX  
PF 15-OCT-1999; 99WO-EP07832.  
XX  
PR 16-OCT-1998; 98EP-0119583.  
XX  
PA (BOEH ) BOEHRINGER INGELHEIM INT GMBH.  
PA (UYHO-) UNIV HOSPITAL LEIDEN.  
XX  
PI Schrier PI, Aarnoudse CA, Heider K, Klade C;  
XX  
XX WPI; 2000-339685/29.  
DR N-PSDB; AAD00152.  
XX  
XX Tumor-associated antigen useful for cancer immunotherapy is encoded by  
PT the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA -  
PT  
PS Example 3; Page 65; 73pp; English.  
XX  
XX The present sequence is the human NY-ESO-1 short variant protein, a  
CC tumour antigen, identified by screening an esophagus carcinoma cDNA  
CC library. This protein is derived from open reading frame (ORF)-2,  
CC that differs from the CAMEL protein (Cytotoxic T lymphocytes (CTL)-  
CC recognised Antigen on MELANOMA), a tumour-associated antigen, by only its  
CC last 5 amino acids. It contains epitopes of tumour specific T-cells.  
CC NY-ESO-1 is expressed in different tumour types, but not in healthy  
CC tissues except in testis. The tumour-associated antigen displayed on  
CC melanoma cells is recognised by cytotoxic T lymphocytes. This sequence  
CC has anticancer activity. CAMEL tumour antigen and immunogenic peptides  
CC derived from it are useful for cancer immunotherapy. They have the

CC potential to induce an immune response, by eliciting a CTL response.  
CC The DNA molecule is used for the construction of recombinant or fusion  
CC proteins.  
XX  
SQ Sequence 58 AA;  
Query Match 100.0%; Score 45; DB 21; Length 58;  
Best Local Similarity 100.0%; Pred. No. 0.13;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMAQALAF 10  
Db 2 LMAQALAF 11  
|||||

RESULT 7  
AAU84819  
ID AAU84819 standard; Protein: 58 AA.  
XX  
AC AAU84819;  
XX  
DT 08-MAY-2002 (first entry)  
XX  
DE Human NYSOLB consensus sequence.  
XX  
KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;  
KW viral infection; human immunodeficiency virus; melanoma;  
KW bacterial infection; Salmonella; Legionella; parasitic infection;  
KW Trypanosoma; Toxoplasma; Giardia.  
XX  
OS Homo sapiens.  
XX  
PN WO200190197-A1.  
XX  
PD 29-NOV-2001.  
XX  
PF 25-MAY-2001; 2001WO-AU00622.  
XX  
PR 26-MAY-2000; 2000AU-0007761.  
XX  
PA (AUSU ) UNIV AUSTRALIAN NAT.  
XX  
XX Thomson SA, Ramshaw IA;  
PI WPI; 2002-147575/19.  
XX  
XX New synthetic polypeptides having several different segments of at  
DR least one parent polypeptide linked together differently compared to  
PT the linkage in the parent polypeptide, for inducing immune response  
PT against a pathogen or cancer  
PT  
PS Example 3; Fig 27; 364pp; English.  
XX  
XX The invention relates to a new synthetic polypeptide (I) comprising  
CC several different segments of at least one parent polypeptide linked  
CC together in a different relationship relative to their linkage in the  
CC parent polypeptide to impede, abrogate or otherwise alter at least one  
CC function associated with the parent polypeptide and for inducing an  
CC immune response against a pathogen or cancer. Also included are a  
CC synthetic polynucleotide encoding and a computer system for  
CC designing the synthetic polypeptides. The synthetic polypeptides and  
CC polynucleotides are referred to as a vaccine. The synthetic polypeptide is  
CC useful for modulating immune responses preferably directed against a  
CC pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix,  
CC colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone  
CC liver, oesophagus, brain, testicle, uterus), as potentiating agents.  
CC Compositions comprising the polypeptide may be used in the treatment or  
CC prophylaxis against viral (such as infections caused by HIV (human  
CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis  
CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial  
CC (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,  
CC Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic  
CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,

CC Trypanosoma, Toxoplasma and Giardia) infections. The present  
 CC sequence is a consensus sequence for a parent protein used to design a  
 CC savine of the invention.

XX SQ Sequence 58 AA;

Query Match 100.0%; Score 45; DB 23; Length 58;  
 Best Local Similarity 100.0%; Pred. No. 0.13;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMAQALAF 10  
 |||||  
 DB 2 LMAQALAF 11

## RESULT 8

AAY70854

ID AAY70854 standard; Protein; 109 AA.

XX AC AAY70854;

XX DT 31-JUL-2000 (first entry)

XX DE Human CTL-recognised Antigen on MELanoma (CAMEL) protein.

XX KW CAMEL; CTL-recognised Antigen on MELanoma; cytotoxic T lymphocyte; CTL;  
 KW tumour-associated antigen; LAGE-1; NY-ESO-1; anticancer; melanoma;  
 KW cancer; immunotherapy; immunogenic peptide; immune response.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT Peptide 1..11

XX FT /label= CTL epitope

XX FT /note= "Immunogenic peptide with ability to elicit a  
 CTL response"

XX FT Peptide 2..10

XX FT /label= Immunogenic peptide

XX FT /note= "Specific for HLA-A3"

XX FT Peptide 2..11

XX FT /label= CTL epitope

XX FT /note= "Immunogenic peptide with ability to elicit a  
 CTL response"

XX FT Peptide 10..18

XX FT /label= CAMEL10

XX FT /note= "Specific for HLA-A2"

XX FT Peptide 16..25

XX FT /label= CAMEL16

XX FT /note= "Specific for HLA-A2"

XX FT Peptide 17..25

XX FT /label= CAMEL17

XX FT /note= "Specific for HLA-A2"

XX FT Peptide 51..59

XX FT /label= Immunogenic peptide

XX FT /note= "Specific for HLA-A3101"

XX FT Peptide 101..109

XX FT /label= Immunogenic peptide

XX FT WO200023584-A1.

XX PD 27-APR-2000.

XX PF 15-OCT-1999; 99WO-EP07832.

XX PR 16-OCT-1998; 98EP-0119583.

XX XX (BOH) BOEHRINGER INGELHEIM INT GMBH.

XX PA (UYHO-) UNIV HOSPITAL LEIDEN.

XX PI Schrier PI, Aarnoudse CA, Heider K, Klade C;

XX DR WPI; 2000-339685/29.

XX DR N-PSDB; AAD00149.

XX PT Tumor-associated antigen useful for cancer immunotherapy is encoded by  
 the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA -

XX PS Claim 1; Page 55; 73pp; English.

XX CC The present protein sequence is the human tumour-associated antigen CAMEL  
 CC (Cytotoxic T lymphocytes (CTL)-recognised Antigen on MELanoma). CAMEL  
 CC protein is encoded by the LAGE-1 gene, a tumour-specific antigen. It is  
 CC different from the LAGE-1 protein, since it is translated from a  
 CC different open reading frame (ORF-1). It shows strong homology with  
 CC NY-ESO-1, a melanoma specific tumour antigen. The tumour-associated  
 CC antigen displayed on melanoma cells is recognised by cytotoxic T  
 CC lymphocytes. CAMEL is expressed in tumour cell lines, tumour tissues  
 CC (e.g. breast and lung) and in restricted number of healthy tissues. This  
 CC sequence has anticancer activity. CAMEL tumour antigen and immunogenic  
 CC peptides derived from it are useful for cancer immunotherapy. They have  
 CC the potential to induce an immune response, by eliciting a CTL response.  
 CC The DNA molecule is used to construct recombinant or fusion proteins.

XX SQ Sequence 109 AA;

Query Match 100.0%; Score 45; DB 21; Length 109;  
 Best Local Similarity 100.0%; Pred. No. 0.24;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMAQALAF 10

|||||

DB 2 LMAQALAF 11

## RESULT 9

AAU85130

ID AAU85130 standard; Protein; 3541 AA.

XX AC AAU85130;

XX DT 08-MAY-2002 (first entry)

XX DE Human melanoma specific savine.

XX KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;  
 KW viral infection; human immunodeficiency virus; melanoma;  
 KW bacterial infection; Salmonella; Legionella; parasitic infection;  
 KW Trypanosoma; Toxoplasma; Giardia.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO200190197-A1.

XX PD 29-NOV-2001.

XX PF 25-MAY-2001; 2001WO-AU00622.

XX PR 26-MAY-2000; 2000AU-0007761.

XX PA (AUSU) UNIV AUSTRALIAN NAT.

XX PI Thomson SA, Ramshaw IA;

XX DR WPI; 2002-147575/19.

XX DR N-PSDB; ABR36950.

XX PT New synthetic polypeptides having several different segments of at  
 least one parent polypeptide linked together differently compared to  
 the linkage in the parent polypeptide, for inducing immune response  
 against a pathogen or cancer

XX PS Example 3; Fig 27; 364pp; English.

XX CC The invention relates to a new synthetic polypeptide (I) comprising  
 CC several different segments of at least one parent polypeptide linked

CC together in a different relationship relative to their linkage in the  
 CC parent polypeptide to impede, abrogate or otherwise alter at least one  
 CC function associated with the parent polypeptide and for inducing an  
 CC immune response against a pathogen or cancer. Also included are a  
 CC synthetic polynucleotide encoding and a computer system for  
 CC designing the synthetic polypeptides. The synthetic polypeptides and  
 CC polynucleotides are referred to as a Savine. The synthetic polypeptide is  
 CC useful for modulating immune responses preferably directed against a  
 CC pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix,  
 CC colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone  
 CC liver, oesophagus, brain, testicle, uterus), as potentiating agents.  
 CC Compositions comprising the polypeptide may be used in the treatment or  
 CC prophylaxis against viral (such as infections caused by HIV (human  
 CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis  
 CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial  
 CC (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,  
 CC Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic  
 CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,  
 CC Trypanosoma, Toxoplasma and Giardia) infections. The present  
 CC sequence is a savine protein of the invention.

XX Sequence 3541 AA;  
 SQ Query Match 100.0%; Score 45; DB 23; Length 3541;  
 Best Local Similarity 100.0%; Pred. No. 8.3;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMAQEALAF 10  
 Db 1755 LMAQEALAF 1764  
 |||||

RESULT 10  
 ABG79072  
 ID ABG79072 standard; Peptide: 10 AA.

XX AC ABG79072;  
 XX DT 15-NOV-2002 (first entry)

XX DE Human CAMEL class I HLA widely expressed antigen peptide #1.

XX Cell penetrating peptide; cancer; tumour; melanoma; thymoma; antigen;  
 KW lymphoma; sarcoma; lung cancer; non-Hodgkin's lymphoma; leukaemia;  
 KW Hodgkin's lymphoma; uterine cancer; cervical cancer; bladder cancer;  
 KW kidney cancer; adenocarcinoma; breast cancer; prostate cancer;  
 KW ovarian cancer; pancreatic cancer; epitope; vaccine; dendritic cell;  
 KW tumour infiltrating lymphocyte; TIL; human leukocyte antigen; HLA;  
 KW cytostatic; human.

XX OS Homo sapiens.  
 XX PN WO200264057-A2.

XX PD 22-AUG-2002.

XX PF 15-FEB-2002; 2002WO-US05212.

XX PR 15-FEB-2001; 2001US-268687P.

XX PA (BAYU ) BAYLOR COLLEGE MEDICINE.

XX PI Wang R;

XX DR WPI; 2002-627577/67.

XX Novel composition for treating a disease in an animal, comprises an  
 PT immune effector cell and cell penetrating peptide associated with an  
 PT antigen or antibody -

XX PS Disclosure; Page 17; 61pp; English.

XX The invention relates to a composition (I) comprising an immune effector

CC cell and a cell penetrating peptide (CPP) associated with an antigen or  
 CC antibody. Also included are (1) a vaccine comprising (I), CPP  
 CC associated with an antigen, and a pharmaceutically acceptable carrier  
 CC and (2) preparing a composition for a disease, by providing (I)  
 CC and CPP associated with an antigen for disease, and introducing the  
 CC antigen-associated CPP to (I), where antigen enters into the cell.  
 CC The antigens are, for example, tumour antigen derived epitopes  
 CC recognised by tumour infiltrating lymphocytes (TIL) of HLA (human  
 CC leukocyte antigen) class I or II. The composition is useful for enhancing  
 CC immunity in an animal to a disease, by administering a mature dendritic  
 CC cell comprising CPP associated with an antigen to disease, to the animal,  
 CC such that following the administration, animal is protected from disease,  
 CC where the animal comprises both CD4+ and CD8+ T cells. It is also useful  
 CC for treating a disease (e.g. cancer, tumour, melanoma, thymoma,  
 CC lymphoma, sarcoma, lung cancer, non-Hodgkin's lymphoma, leukaemia,  
 CC Hodgkin's lymphoma, uterine cancer, cervical cancer, bladder cancer,  
 CC kidney cancer, adenocarcinoma, breast cancer, prostate cancer,  
 CC ovarian cancer, and pancreatic cancer). The animal is further subjected to  
 CC a cancer treatment including surgery, radiation, chemotherapy or gene  
 CC therapy. The administration of (I), preferably dendritic cell is prior  
 CC to, subsequent to or concurrent with, the cancer treatment. The present  
 CC sequence is a tumour antigen derived epitope for inclusion in the  
 CC composition of the invention.

XX Sequence 10 AA;

Query Match 91.1%; Score 41; DB 23; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.13;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMAQEALAF 9  
 Db 2 LMAQEALAF 10  
 |||||

RESULT 11

ABP73965  
 ID ABP73965 standard; Protein; 222 AA.

XX AC ABP73965;

XX DT 30-JAN-2003 (first entry)

XX Candida albicans essential protein SEQ ID NO 7802.

XX Fungus; yeast; tetracycline; promoter; GRACE strain; biosynthesis;  
 KW signal transduction; DNA replication; cell division; growth;  
 KW proliferation; Candida albicans; fungicide; antifungal.

XX OS Candida albicans.

XX PN WO200253728-A2.

XX PD 11-JUL-2002.

XX PF 26-DEC-2001; 2001WO-US49486.

XX PR 29-DEC-2000; 2000US-259128P.

XX PR 20-FEB-2001; 2001US-0792024.

XX PR 22-AUG-2001; 2001US-314050P.

XX PA (ELIT-) ELITRA PHARM INC.

XX Roemer T, Jiang B, Boone C, Bussey H, Ohlsen KL;

XX WPI; 2002-566694/60.

XX DR N-PSDB; ABZ32515.

XX Constructing strains for identifying gene products as effective targets  
 PT for therapeutic intervention, by inactivating in the strain one allele  
 PT of a gene and placing other allele of the gene under conditional  
 PT expression -

XX

PS Claim 44; SEQ ID NO 7802; 167pp + Sequence Listing; English.

XX The invention relates to constructing (M1) a strain of diploid fungal cells in which both alleles of a gene are modified, comprising modifying one allele by insertion or replacement by a cassette having an expressible selectable marker and modifying other allele by recombination, of a promoter replacement fragment with a heterologous promoter, so that expression of the second allele is regulated by the promoter. (M1) is useful for constructing a strain of diploid fungal cells in which both alleles of a gene are modified. The diploid fungal cells having both alleles modified are useful for identifying a gene that is essential to the survival or growth of a fungus, a gene that contributes to the virulence and/or pathogenicity of a fungus, a gene that contributes to the resistance of a diploid fungus to an antifungal agent, an antifungal agent that inhibits the growth of a diploid fungus and for identifying a therapeutic agent for treatment of a mammalian disease. (M1) is useful for identifying a compound which modulates the activity of a gene product, preferably enzymatic activity, carbon compound catabolism, biosynthetic, transporter, transcriptional, translational, signal transduction, DNA replication and cell division activity. The method is useful for identifying a compound having the ability to inhibit growth or proliferation of C. albicans cells and for treating infection by C. albicans. The present sequence is that of an essential Candida albicans protein used in the method of the invention. CC Note: The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied to Derwent by the European Patent Office.

XX Sequence 222 AA;

Query Match 75.6%; Score 34; DB 23; Length 222;

Best Local Similarity 80.0%; Pred. NO. 73;

Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMAQEALAFLL 10

Db 122 LMAQEADAFV 131

RESULT 12

ABU65209

ID ABU65209 standard; Protein; 577 AA.

XX AC ABU65209;

XX 20-MAY-2003 (first entry)

XX Human NOV12la protein.

XX NOVX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;

KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;

KW human.

XX Homo sapiens.

XX WO200272757-A2.

XX 19-SEP-2002.

PF 08-MAR-2002; 2002WO-US06908.

XX 08-MAR-2001; 2001US-274101P.

PR 08-MAR-2001; 2001US-274194P.

PR 08-MAR-2001; 2001US-274281P.

PR 08-MAR-2001; 2001US-274322P.

PR 09-MAR-2001; 2001US-274849P.

PR 12-MAR-2001; 2001US-275235P.

PR 13-MAR-2001; 2001US-275578P.

PR 13-MAR-2001; 2001US-275578P.

PR 13-MAR-2001; 2001US-275601P.

PR 14-MAR-2001; 2001US-276000P.

PR 16-MAR-2001; 2001US-276776P.

PR 19-MAR-2001; 2001US-276994P.

PR 20-MAR-2001; 2001US-277239P.

PR 20-MAR-2001; 2001US-277321P.

PR 20-MAR-2001; 2001US-277327P.

PR 21-MAR-2001; 2001US-277791P.

PR 22-MAR-2001; 2001US-277833P.

PR 23-MAR-2001; 2001US-278152P.

PR 26-MAR-2001; 2001US-278894P.

PR 27-MAR-2001; 2001US-278999P.

PR 27-MAR-2001; 2001US-279036P.

PR 28-MAR-2001; 2001US-279344P.

PR 30-MAR-2001; 2001US-277338P.

PR 30-MAR-2001; 2001US-279955P.

PR 30-MAR-2001; 2001US-280233P.

PR 02-APR-2001; 2001US-280802P.

PR 02-APR-2001; 2001US-280822P.

PR 02-APR-2001; 2001US-280900P.

PR 04-APR-2001; 2001US-281194P.

PR 13-APR-2001; 2001US-283675P.

PR 30-APR-2001; 2001US-287424P.

PR 02-MAY-2001; 2001US-288066P.

PR 03-MAY-2001; 2001US-288342P.

PR 03-MAY-2001; 2001US-288528P.

PR 15-MAY-2001; 2001US-291190P.

PR 16-MAY-2001; 2001US-291099P.

PR 16-MAY-2001; 2001US-291240P.

PR 30-MAY-2001; 2001US-294485P.

PR 31-MAY-2001; 2001US-294889P.

PR 31-MAY-2001; 2001US-294899P.

PR 18-JUN-2001; 2001US-299027P.

PR 19-JUN-2001; 2001US-299303P.

PR 19-JUN-2001; 2001US-299310P.

PR 10-JUL-2001; 2001US-304354P.

PR 31-JUL-2001; 2001US-309198P.

PR 16-AUG-2001; 2001US-312903P.

PR 10-SEP-2001; 2001US-318462P.

PR 12-SEP-2001; 2001US-318770P.

PR 27-SEP-2001; 2001US-325430P.

PR 27-SEP-2001; 2001US-325681P.

PR 18-OCT-2001; 2001US-330380P.

PR 31-OCT-2001; 2001US-335301P.

PR 14-NOV-2001; 2001US-332172P.

PR 14-NOV-2001; 2001US-332271P.

PR 14-NOV-2001; 2001US-332272P.

PR 14-NOV-2001; 2001US-333184P.

PR 14-NOV-2001; 2001US-333272P.

PR 21-NOV-2001; 2001US-332094P.

PR 03-DEC-2001; 2001US-337426P.

PR 03-DEC-2001; 2001US-338092P.

PR 04-DEC-2001; 2001US-337185P.

PR 03-JAN-2002; 2002US-345705P.

PR 07-MAR-2002; 2002US-0092900.

(CURA-) CURAGEN CORP.

Padiqaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;  
Zerhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;  
Patturajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;  
Fernandes ER, Casman SJ, Malyankar DM, Gerlach V, Liu Y;  
Anderson D, Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H;  
Alsobrook JP, Lepley DM, Rieger DK;

WPI: 2002-723332/78.

N-PSDB; ABX97176.

NOVX polypeptides and polynucleotides, useful for preventing or treating a disorder associated with aberrant NOVX expression or activity e.g., cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial asthma

Claim 1; Page 455; 1103pp; English.

This invention describes novel human NOVX polypeptides which have cytostatic, cardiant, antiarteriosclerotic, antiasthmatic and

Tue Oct 7 15:52:23 2003

```
CC hypotensive activity. Pharmaceutical compositions comprising the NOVX
CC proteins or nucleic acid molecules or NOVX antibodies are useful for
CC preventing or treating a disorder associated with aberrant NOVX
CC expression or activity e.g. cancer, hypertension, atherosclerosis,
CC cardiomyopathy or bronchial asthma. The products of the invention can
CC be used for gene therapy or in a vaccine. ABU65041-ABU65218 represent
CC the NOVX polypeptides encoded by ABX97008-ABX97185.
XX
XX
SQ Sequence 577 AA;
Query Match 75.6%; Score 34; DB 23; Length 577;
Best Local Similarity 70.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 LMAOEALAF 10
II :||||:
Db 260 LMLKEALAFI 269
RESULT 13
AAG22016
ID AAG22016 standard; Protein; 320 AA.
XX
XX
AC AAG22016;
DT 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 24785.
XX
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX
OS Arabidopsis thaliana.
XX
XX
PN EP1033405-A2.
XX
XX
PD 06-SEP-2000.
XX
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0123788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
```

PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 06-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 14-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.

PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.  
Query Match 73.3%; Score 33; DB 21; Length 320;  
Best Local Similarity 77.8%; Pred. No. 1.7e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 2 MAQEALAF 10  
Db 1 MQQEALSF 9  
RESULT 14  
AAB47134  
ID AAB47134 standard; Protein; 761 AA.  
XX  
AC AAB47134;  
XX  
DT 04-JUN-2001 (first entry)  
XX  
DE CDIFF-15, Incyte ID No. 3478571CD1.  
XX  
KW Human; cell differentiation; CDIFF; agonist; antagonist; epilepsy;  
cell proliferation; Alzheimer's disease; schizophrenic disorder;  
KW arteriosclerosis; cancer; atherosclerosis; diabetes mellitus.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 2  
FT Modified-site /note= "Potential phosphorylation site"  
FT Modified-site 67  
FT Modified-site /note= "Potential phosphorylation site"  
FT Modified-site 84  
FT Modified-site /note= "Potential phosphorylation site"  
FT Modified-site 123  
FT Modified-site /note= "Potential phosphorylation site"  
FT Modified-site 148  
FT Modified-site /note= "Potential glycosylation site"  
FT Modified-site 169  
FT Modified-site /note= "Potential glycosylation site"  
FT Modified-site 193  
FT Modified-site /note= "Potential phosphorylation site"  
FT Modified-site 250  
FT Modified-site /note= "Potential phosphorylation site"  
FT Modified-site 269  
FT Modified-site /note= "Potential phosphorylation site"  
FT Modified-site 355  
FT Modified-site /note= "Potential phosphorylation site"  
FT Modified-site 382  
FT Modified-site /note= "Potential phosphorylation site"  
FT Modified-site 457  
FT Modified-site /note= "Potential phosphorylation site"  
FT Modified-site 473  
FT Modified-site /note= "Potential phosphorylation site"  
FT Modified-site 485  
FT Modified-site /note= "Potential phosphorylation site"  
FT Modified-site 499  
FT Modified-site /note= "Potential phosphorylation site"  
FT Modified-site 557  
FT Modified-site /note= "Potential phosphorylation site"  
FT Modified-site 582  
FT Modified-site /note= "Potential phosphorylation site"  
FT Modified-site 622  
FT Modified-site /note= "Potential phosphorylation site"  
FT Modified-site 631  
FT Modified-site /note= "Potential phosphorylation site"  
FT Modified-site 662  
FT Modified-site /note= "Potential phosphorylation site"  
FT Modified-site 671  
FT Modified-site /note= "Potential phosphorylation site"  
XX WO200119860-A2.

Tue Oct 7 15:52:23 2003

XX PD 22-MAR-2001.  
XX PF 14-SEP-2000; 2000WO-US25435.  
XX PR 15-SEP-1999; 99US-0154140.  
XX PR 06-DEC-1999; 99US-0169155.  
XX PA (INCY-) INCYTE GENOMICS INC.  
XX XX Tang YT, Hillman JL, Yue H, Reddy R, Lal P, Shah P, Azimzai Y;  
XX PI Baughn MR, Lu DAM, Bandman O, Shih LL, Patterson C;  
XX DR N-PSDB; AAC85556.  
XX DR WPI: 2001-211447/21.  
XX DR N-PSDB; AAC85556.  
XX PT Isolated polypeptides and polynucleotides involved in cell  
XX PT differentiation are used for treatment, prevention and diagnosis of  
XX PT cell proliferative, developmental and neurological disorders e.g.  
XX PT cancer and Alzheimer's disease -  
XX PS Claim 1; Page 107-109; 137pp; English.  
XX CC The sequences given in AAB47123-50 show human polypeptides involved  
XX CC in cell differentiation (CDIFF). CDIFF polypeptides and agonists of  
XX CC these are used to treat a disease or condition associated with  
XX CC decreased expression of functional CDIFF. An antagonist of CDIFF is  
XX CC used to treat a disease or condition associated with over expression  
XX CC of functional CDIFF. CDIFF polypeptides may be used for the treatment,  
XX CC prevention and diagnosis of cell proliferative, developmental and  
XX CC neurological disorders, such as Alzheimer's disease, schizophrenic  
XX CC disorders, arteriosclerosis, cancer, atherosclerosis, diabetes mellitus  
XX CC and epilepsy. The CDIFF-15 sequence is homologous to Mus musculus  
XX CC schlafen3.  
XX SQ Sequence 761 AA;  
Query Match 73.3%; Score 33; DB 22; Length 761;  
Best Local Similarity 87.5%; Pred. No. 4e+02;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 3 AQEALAF 10  
Db 130 SQEALAF 137  
RESULT 15  
AAE22231  
ID AAE22231 standard; Protein; 891 AA.  
XX AC AAE22231;  
XX XX 25-JUL-2002 (first entry)  
XX DE Human schlafen C protein.  
XX XX Human; morphogenesis; DNAX interferon like receptor subunit 4; DIRS4;  
XX KW immune system; cytokine receptor; tumour necrosis factor; TNF; TNF;  
XX KW toll like receptor like molecule; TLR-L; transforming growth factor;  
XX KW TGF; 5685C6; claudin; schlafen C; gene therapy; vaccine; immunological;  
XX KW medical disorder.  
XX OS Homo sapiens.  
XX XX WO200220569-A2.  
XX XX 14-MAR-2002.  
XX PD 07-SEP-2001; 2001WO-US28013.  
XX PF 08-SEP-2000; 2000US-231267P.  
XX PR (SCHE ) SCHERING CORP.  
XX PA

XX PI Parham CL, Gorman DM, Kurata H, Arai N, Sana TR, Mattson JD;  
XX PI Murphy EE, Savkoor C, Grein J, Smith KM, Mcclanahan TK;  
XX DR WPI: 2002-362239/39.  
XX DR N-PSDB; AAD35343.  
XX XX Recombinant polypeptide for immunizing a subject, comprises  
XX PT non-overlapping segments of amino acids identical to cytokine receptor  
XX PT sequences -  
XX PS Claim 1; Page 187-191; 211pp; English.  
XX CC The present invention relates to compositions and methods for affecting  
XX CC mammalian physiology, including morphogenesis or immune system function.  
XX CC The invention particularly relates to recombinant polypeptides comprising  
XX CC 3 distinct non-overlapping segments of four amino acids identical to  
XX CC cytokine receptors, e.g., cytokine receptor like molecular structures  
XX CC such as segments of the sequences of DNAX interferon like receptor  
XX CC subunit 4 (DIRS4), tumour necrosis factor (TNF or TNF), toll like  
XX CC receptor like molecules (TLR-L1 through TLR-L5), transforming growth  
XX CC factor (TGF), 5685C6, claudins and schlafens. Sequences of the invention  
XX CC are used to modulate physiology or development of a cell. They are also  
XX CC used in gene therapy and as vaccines. Nucleic acid sequences are useful  
XX CC as probes for detecting a level of respective genes or transcripts in  
XX CC patients suspected of having an immunological or other medical disorders.  
XX CC The present sequence is human schlafen C protein.  
XX SQ Sequence 891 AA;  
Query Match 73.3%; Score 33; DB 23; Length 891;  
Best Local Similarity 87.5%; Pred. No. 4.7e+02;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 3 AQEALAF 10  
Db 130 SQEALAF 137  
Search completed: October 7, 2003, 13:25:30  
Job time : 17.5823 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:19:12 ; Search time 5.63291 Seconds  
(without alignments)  
170.726 Million cell updates/sec

Title: US-09-807-512-12  
Perfect score: 45  
Sequence: 1 LMAQEALAF 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues  
Total number of hits satisfying chosen parameters: 2833308

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 10%
                  Listing first 45

```

```
Database :      PIR_76:*
1:  pir1:*
2:  pir2:*
3:  pir3:*
4:  pir4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	33	73.3	207	2	H71671	ribosomal protein	
2	33	73.3	333	2	F84775	hypothetical prote	
3	32	71.1	96	1	I64010	C4-dicarboxylate t	
4	32	71.1	145	2	D84131	PTS system, mannit	
5	32	71.1	205	2	G82358	conserved hypothet	
6	32	71.1	215	2	T4B554	lysine decarboxyla	
7	32	71.1	2697	2	T25444	hypothetical prote	
8	31	68.9	398	2	F69266	3-ketoacyl-CoA thi	
9	31	68.9	439	1	VHBpCL	minor capsid prote	
10	31	68.9	636	2	G86504	signal peptidase I	
11	31	68.9	636	2	F72118	signal peptidase I	
12	31	68.9	922	2	S37871	hypothetical prote	
13	31	68.9	1225	2	A49464	chromosome segrega	
14	31	68.9	4131	2	T21085	hypothetical prote	
15	30	66.7	105	2	AB1360	protein gp21 [Bact	
16	30	66.7	238	2	AE0918	conserved hypothet	
17	30	66.7	263	2	AE0631	2,4-dihydroxyhept-	
18	30	66.7	274	2	D95165	NH(3)-dependent NA	
19	30	66.7	274	2	C98031	NAD synthase (glut	
20	30	66.7	275	1	PNEBT	beta-lactamase (EC	
21	30	66.7	308	2	D83629	probable permease	
22	30	66.7	363	2	E86878	DNA-damage-inducib	
23	30	66.7	424	2	B83928	multiple sugar tra	
24	30	66.7	437	2	B70353	conserved hypothet	
25	30	66.7	454	2	T16825	hypothetical prote	
26	30	66.7	498	2	T07384	probable hexokinase	
27	30	66.7	564	2	S29246	uracinate hydratase	
28	29	64.4	81	2	T22176	hypothetical prote	
29	29	64.4	81	2	S26842	transcription requ	

## ALIGNMENTS

RESULT 1  
H71671

Ribosomal protein L4 - *Rickettsia prowazekii*  
C:Species: *Rickettsia prowazekii*  
C:Date: 21-Nov-1998 #sequence\_revision 21-Nov-1998 #text\_change 03-Nov-2000  
C:Accession: H71671  
R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.;  
Nature 395, 133-140, 1998  
A:Title: The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria  
A:Reference number: A71630; MUID:99039499; PMID:9823893  
A:Accession: H71671  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-207 <AND>  
A:Cross-references: GB:AJ235272; GB:AJ235269; NID:g3861033; PIDN:CAAL15098.1; I:  
A:Experimental source: strain Madrid E  
C:Genetics:  
A:Gene: rplD; RP658  
C:Superfamily: Escherichia coli ribosomal protein L4

Query Match	73.3%	Score 33;	DB 2;	Length 207;
Best Local Similarity	60.0%;	Pred.	No. 15;	
Matches	6;	Conservative	4;	Mismatches
				Gaps 0;
				Indels 0;

```
Qy      1 LMAQEALAF 10
         |::|::|
Db     193 LLSOEAVSFL 202
```

RESULT 2  
F84775

F04773  
hypothetical protein Atg36000 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C:Accession: F84775  
R:Llin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii,  
M.; Koo, H.; Moffat, K.S.; Gronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tal  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Ve  
Nature 402, 761-768, 1999  
A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: F84775  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-333 <SFO>  
A:Cross-references: GB:AE002093; NID:g4510369; PIDN:AAD21457.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: Atg36000  
A:Map position: 2

	Query Match	Score	DB	Length
Best Local Similarity	73.3%	33	2	333
Pred. No.	77.8%	25		

hypothetical protein  
DNA-binding response protein  
probable ornithine decarboxylase  
uridine 5'-monophosphate decarboxylase  
ornithine-5'-phosphotransferase  
hypothetical protein  
uncharacterized protein  
NH3-dependent NAD-dependent  
NAD+ synthase (glutamine synthetase)  
hypothetical protein  
probable oligopeptidase  
hypothetical protein  
hypothetical protein  
hypothetical protein  
probable transcription factor  
oxidoreductase (NADH-dependent)

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.  
A:Reference number: A82035; MUID:20406833; PMID:10952301  
A:Accession: G82358  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-205 <HEI>  
A:Cross-references: GB:AE004104; GB:AE003852; NID:g9654534; PIDN:AAF93313.1; GSPDB:GN  
A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
C:Genetics:  
A:Gene: VC0136  
A:Map position: 1  
C:Superfamily: hypothetical protein bl798

Query Match 71.1%; Score 32; DB 2; Length 205;  
Best Local Similarity 77.8%; Pred. No. 25;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMAQEALAF 9  
|:|:|:|:|:|  
DB 64 LVAQSALAF 72

RESULT 6  
T48554  
Lysine decarboxylase-like protein - Arabidopsis thaliana  
N:Alternate names: protein F14F18.120  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 17-Nov-2000  
C:Accession: T48554  
R:Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft  
submitted to the Protein Sequence Database, April 2000  
A:Reference number: Z24490  
A:Accession: T48554  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-215 <BEV>  
A:Cross-references: EMBL:AL163812  
A:Experimental source: Cultivar Columbia; BAC clone F14F18  
C:Genetics:  
A:Map position: 5  
A:Insertions: 37/3; 70/2; 80/3; 114/1; 136/3; 184/3  
A:Note: F14F18.120  
C:Superfamily: yeast conserved hypothetical protein YJL055W

Query Match 71.1%; Score 32; DB 2; Length 215;  
Best Local Similarity 77.8%; Pred. No. 26;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 MAQEALAF 10  
|:|:|:|:|:|  
DB 102 MAQEAEAFI 110

RESULT 7  
T25444  
Hypothetical protein B0261.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jan-2000  
C:Accession: T25444  
R:Scheet, P.; Gattung, S.  
submitted to the EMBL Data Library, April 1997  
A:Description: The sequence of C. elegans cosmid B0261.  
A:Reference number: Z20036  
A:Accession: T25444  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-2697 <SCH>  
A:Cross-references: EMBL:U97016; PIDN:AA52354.1; GSPDB:GN00019; CESP:B0261.2  
A:Experimental source: strain Bristol N2; clone B0261  
C:Genetics:  
A:Gene: CESP:B0261.2

QY 3 AQEALAF 10  
|:|:|:|:|:|  
DB 133 AQEVLAF 140

RESULT 5  
G82358  
conserved hypothetical protein VC0136 [imported] - *Vibrio cholerae* (strain N16961 serogr  
C:Species: *Vibrio cholerae*  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
A:Reference number: A64000; MUID:95350630; PMID:7542800  
A:Accession: I64010  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-96 <TIGR>  
A:Cross-references: GB:U32746; GB:L42023; NID:gl573626; PIDN:AAC22295.1; PID:gl573632;  
C:Superfamily: C4-dicarboxylate carrier protein

Query Match 71.1%; Score 32; DB 1; Length 96;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AQEALAF 9  
|:|:|:|:|:|  
DB 42 AQEALAF 48

RESULT 4  
D84131  
PTS system, mannitol-specific enzyme II, A component BH3852 [imported] - *Bacillus halodurans*  
C:Species: *Bacillus halodurans*  
C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
C:Accession: D84131  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and  
A:Reference number: A83650; MUID:20512582; PMID:11058132  
A:Accession: D84131  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-145 <STO>  
A:Cross-references: GB:AP001520; GB:BA000004; NID:gl0176401; PIDN:BA807571.1; GSPDB:GN00  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: BH3852  
C:Superfamily: phosphotransferase system mannitol-specific enzyme II factor III; phospho

Query Match 71.1%; Score 32; DB 2; Length 145;  
Best Local Similarity 87.5%; Pred. No. 18;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 AQEALAF 10  
|:|:|:|:|:|  
DB 133 AQEVLAF 140

RESULT 5  
G82358  
conserved hypothetical protein VC0136 [imported] - *Vibrio cholerae* (strain N16961 serogr  
C:Species: *Vibrio cholerae*  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
A:Reference number: A64000; MUID:95350630; PMID:7542800  
A:Accession: I64010  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-96 <TIGR>  
A:Cross-references: GB:U32746; GB:L42023; NID:gl573626; PIDN:AAC22295.1; PID:gl573632;  
C:Superfamily: C4-dicarboxylate carrier protein

Query Match 71.1%; Score 32; DB 1; Length 96;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AQEALAF 9  
|:|:|:|:|:|  
DB 42 AQEALAF 48

RESULT 4  
D84131  
PTS system, mannitol-specific enzyme II, A component BH3852 [imported] - *Bacillus halodurans*  
C:Species: *Bacillus halodurans*  
C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
C:Accession: D84131  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and  
A:Reference number: A83650; MUID:20512582; PMID:11058132  
A:Accession: D84131  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-145 <STO>  
A:Cross-references: GB:AP001520; GB:BA000004; NID:gl0176401; PIDN:BA807571.1; GSPDB:GN00  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: BH3852  
C:Superfamily: phosphotransferase system mannitol-specific enzyme II factor III; phospho

Query Match 71.1%; Score 32; DB 2; Length 145;  
Best Local Similarity 87.5%; Pred. No. 18;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 AQEALAF 10  
|:|:|:|:|:|  
DB 133 AQEVLAF 140

RESULT 5  
G82358  
conserved hypothetical protein VC0136 [imported] - *Vibrio cholerae* (strain N16961 serogr  
C:Species: *Vibrio cholerae*  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
A:Reference number: A64000; MUID:95350630; PMID:7542800  
A:Accession: I64010  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-96 <TIGR>  
A:Cross-references: GB:U32746; GB:L42023; NID:gl573626; PIDN:AAC22295.1; PID:gl573632;  
C:Superfamily: C4-dicarboxylate carrier protein

Query Match 71.1%; Score 32; DB 1; Length 96;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AQEALAF 9  
|:|:~:~:~:~:~:~  
DB 42 AQEALAF 48

RESULT 4  
D84131  
PTS system, mannitol-specific enzyme II, A component BH3852 [imported] - *Bacillus halodurans*  
C:Species: *Bacillus halodurans*  
C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
C:Accession: D84131  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and  
A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Map position: 1  
 A:Introns: 41/3; 75/2; 173/3; 283/3; 354/2; 374/3; 432/2; 470/2; 569/3; 607/3; 667/1; 791/3; 228/3; 2337/3; 2516/2; 2562/3; 2652/1  
 C:Superfamily: Yeast TOR2 protein

Query Match 71.1%; Score 32; DB 2; Length 2697;  
 Best Local Similarity 77.8%; Pred. No. 3.7e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LMAQEALAF 10  
 II :|||||  
 Db 163 MASKALAF 171

## RESULT 8

F69266  
 3-ketoacyl-CoA thiolase (acab-4) homolog - Archaeoglobus fulgidus  
 C:Species: Archaeoglobus fulgidus  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999  
 C:Accession: F69266  
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
 ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.  
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
 Nature 390, 364-370, 1997  
 A:Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.  
 Smith, H.O.; Woese, C.R.; Venter, J.C.  
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo  
 A:Reference number: A69250; MUID:98049343; PMID:9389475  
 A:Accession: F69266  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-398 <KLE>  
 A:Cross-references: GB:AE001097; GB:AE000782; NID:g2689420; PIDN:AAB91095.1; PID:g265051  
 C:Superfamily: probable 3-ketoacyl-CoA thiolase MTH177

Query Match 68.9%; Score 31; DB 2; Length 398;  
 Best Local Similarity 77.8%; Pred. No. 83;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LMAQEALAF 9  
 III |||| I  
 Db 304 LMAEALGF 312

## RESULT 9

VHBPCL  
 minor capsid protein precursor C - phage lambda  
 N:Contains: capsid assembly protein Nu3  
 C:Species: phage lambda  
 C:Date: 13-Jun-1983 #sequence\_revision 13-Jun-1983 #text\_change 23-Jul-1999  
 C:Accession: A04333; B92891  
 R:Daniels, D.  
 submitted to the Nucleic Acid Sequence Database, September 1982  
 A:Reference number: A94614  
 A:Accession: A04333  
 A:Molecule type: DNA  
 A:Residues: 1-439 <DAN>  
 R:Sanger, F.; Coulson, A.R.; Hong, G.F.; Hill, D.F.; Petersen, G.B.  
 J. Mol. Biol. 162, 729-773, 1982  
 A:Title: Nucleotide sequence of bacteriophage lambda DNA.  
 A:Reference number: A92891; MUID:83189071; PMID:6221115  
 A:Accession: B92891  
 A:Molecule type: DNA  
 A:Residues: 1-439 <SAN>  
 A:Cross-references: GB:J02459; GB:M17233; GB:M24325; GB:V00636; GB:X00906; NID:g215104;  
 A:Note: there are five alternative initiation sites for gene C translation; the codon co  
 A:Note: gene Nu3 is believed to be coded within gene C, in the same reading frame; there  
 C:Comment: During head assembly, the gene C protein is covalently linked with an equimol  
 C:Comment: In conjunction with host protein groE, gene C protein specifically cleaves s  
 C:Comment: Gene Nu3 protein forms the scaffold for capsid assembly and is required for h  
 C:Genetics:  
 A:Gene: C  
 A:Map position: 9.11-11.82

C:Superfamily: phage lambda minor capsid protein C  
 F:260-439/Product: capsid assembly protein Nu3 (version 1) #status predicted <CAL>  
 F:309-439/Product: capsid assembly protein Nu3 (version 2) #status predicted <CA2>

Query Match 68.9%; Score 31; DB 1; Length 439;  
 Best Local Similarity 80.0%; Pred. No. 92;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LMAQEALAF 10  
 I ||||| I  
 Db 52 LTAQEALATL 61

## RESULT 10

G86504  
 signal peptidase I [imported] - Chlamydophila pneumoniae (strain J138)  
 C:Species: Chlamydophila pneumoniae, Chlamydia pneumoniae  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 02-Mar-2001  
 C:Accession: G86504  
 R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.;  
 Nucleic Acids Res. 28, 2311-2314, 2000  
 A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.  
 A:Reference number: A86491; MUID:20330349; PMID:10871362  
 A:Accession: G86504  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-636 <STO>  
 A:Cross-references: GB:BA000008; NID:g8978484; PIDN:BA098321.1; GSPDB:GN00142  
 A:Experimental source: strain J138  
 C:Genetics:  
 A:Gene: lepB

Query Match 68.9%; Score 31; DB 2; Length 636;  
 Best Local Similarity 75.0%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 MAQEALAF 9  
 :||:||||  
 Db 58 LAQALAF 65

## RESULT 11

F72118  
 signal peptidase I, probable CP0664 [imported] - Chlamydophila pneumoniae (strains CW  
 C:Species: Chlamydophila pneumoniae, Chlamydia pneumoniae  
 C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 11-May-2000  
 C:Accession: F72118; B81552  
 R:Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood,  
 Nature Genet. 21, 385-389, 1999  
 A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.  
 A:Reference number: A72000; MUID:99206606; PMID:10192388  
 A:Accession: F72118  
 A:Molecule type: DNA  
 A:Residues: 1-636 <ARN>  
 A:Cross-references: GB:AE001597; GB:AE001363; NID:g4376367; PIDN:AAD18263.1; PID:g437  
 A:Experimental source: strain CWL029  
 R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke  
 ; C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe  
 Nucleic Acids Res. 28, 1397-1406, 2000  
 A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39  
 A:Reference number: A81500; MUID:20150255; PMID:10684935  
 A:Accession: B81552  
 A:Molecule type: DNA  
 A:Residues: 1-636 <REA>  
 A:Cross-references: GB:AE002224; GB:AE002161; NID:g7189574; PIDN:AAF38476.1; PID:g718  
 A:Experimental source: strain AR39, HL cells  
 C:Genetics:  
 A:Gene: lepB; CP0664

Query Match 68.9%; Score 31; DB 2; Length 636;  
 Best Local Similarity 75.0%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Tue Oct 7 15:52:25 2003

QY 2 MAQEALAF 9  
 :||:||||  
 Db 58 LAQALAF 65

## RESULT 12

S37871  
 hypothetical protein YKL050c - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: hypothetical protein YKL301  
 C:Species: Saccharomyces cerevisiae  
 A:Variety: strain S288C  
 C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 19-Apr-2002  
 C:Accession: S37871; S37872; S39186; S44531  
 R:Purnelle, B.; Skala, J.; van Dyck, L.; Tettelin, H.; Goffeau, A.  
 submitted to the Protein Sequence Database, March 1994  
 A:Reference number: S37851  
 A:Accession: S37871  
 A:Molecule type: DNA  
 A:Residues: 1-452 <PUR>  
 A:Cross-references: EMBL:Z28050; MIPS:YKL050c  
 R:Rasmussen, S.; von Wettstein, D.  
 submitted to the Protein Sequence Database, March 1994  
 A:Reference number: S37872  
 A:Accession: S37872  
 A:Molecule type: DNA  
 A:Residues: 320-922 <PAS>  
 A:Cross-references: EMBL:Z28050; MIPS:YKL050c  
 R:Rasmussen, S.W.  
 submitted to the EMBL Data Library, November 1993  
 A:Reference number: S39186  
 A:Accession: S39186  
 A:Molecule type: DNA  
 A:Residues: 410-922 <RA2>  
 A:Cross-references: EMBL:X75781; NID:q433634; PIDN:CAA53420.1; PID:q433645  
 R:Rasmussen, S.W.  
 Yeast 10, 63-68, 1994  
 A:Title: Sequence of a 28.6 kb region of yeast chromosome XI includes the FBAl and TOA2  
 otifs also found in plant storage proteins and 13 ORFs with weak or no homology to known  
 A:Reference number: S44521  
 A:Accession: S44531  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 410-922 <RAW>  
 A:Cross-references: EMBL:X75781; NID:q433634; PIDN:CAA53420.1; PID:q433645  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993  
 C:Genetics:  
 A:Cross-references: SGD:S0001533  
 A:Map position: 111

Query Match 68.9%; Score 31; DB 2; Length 922;  
 Best Local Similarity 77.8%; Pred. No. 2e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMAQALAF 9  
 :||:||||  
 Db 173 LAQKALAF 181

## RESULT 13

A49464  
 chromosome segregation protein SMCl - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: probable head-rod-tail protein SMCl; protein YFL008W  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 20-Jun-2000  
 C:Accession: A49464; S56246; S62305; S41804  
 R:Strunnikov, A.V.; Laktionov, V.L.; Koshland, D.  
 J. Cell Biol. 123, 1635-1648, 1993  
 A:Title: SMCl: an essential yeast gene encoding a putative head-rod-tail protein is requ  
 A:Reference number: A49464; MUID:94103320; PMID:8276886  
 A:Accession: A49464  
 A:Molecule type: DNA  
 A:Residues: 1-1225 <STR>  
 A:Cross-references: GB:L00602; NID:g172620; PIDN:AAA16595.1; PID:g172621

R:Murakami, Y.; Naitou, M.; Hagiwara, H.; Shibata, T.; Ozawa, M.; Sasanuma, S.I.; Sas  
 submitted to the EMBL Data Library, May 1995  
 A:Description: Analysis of the nucleotide sequence of chromosome VI from Saccharomyces  
 A:Reference number: S56186  
 A:Accession: S56246  
 A:Molecule type: DNA  
 A:Residues: 1-1225 <MUR>  
 A:Cross-references: EMBL:D50617; NID:g836685; PIDN:BAA09230.1; PID:g836746; MIPS:YFL0  
 R:Murakami, Y.  
 submitted to the EMBL Data Library, April 1994  
 A:Reference number: S62302  
 A:Accession: S62305  
 A:Molecule type: DNA  
 A:Residues: 1-1225 <MUW>  
 A:Cross-references: EMBL:D31600; NID:g836814; PIDN:BAA06496.1; PID:g836816  
 C:Genetics:  
 A:Gene: SGD:SMCl  
 A:Cross-references: SGD:S0001886; MIPS:YFL008W  
 A:Map position: 6L  
 C:Function:  
 A:Description: probably involved in chromosome segregation  
 C:Superfamily: chromosome segregation protein SMCl  
 C:Keywords: cell division; coiled coil; mitosis; nucleus

Query Match 68.9%; Score 31; DB 2; Length 1225;  
 Best Local Similarity 66.7%; Pred. No. 2.7e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 MAQEALAF 10  
 :||:||||  
 Db 566 VAOEALAF 574

## RESULT 14

T21085  
 hypothetical protein F18C12.1 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 17-Mar-2000  
 C:Accession: T21085  
 R:Harris, B.  
 submitted to the EMBL Data Library, June 1996  
 A:Reference number: Z19371  
 A:Accession: T21085  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-4131 <MIL>  
 A:Cross-references: EMBL:Z75536; PIDN:CAA99830.1; GSPDB:GN00019; CESP:F18C12.1  
 A:Experimental source: clone F18C12  
 C:Genetics:  
 A:Gene: CESP:F18C12.1  
 A:Map position: 1  
 A:Introns: 54/3; 112/3; 146/1; 294/1; 381/3; 455/3; 596/3; 716/3; 743/2; 782/3; 820/3  
 09/3; 2303/3; 2483/1; 2529/1; 2612/3; 2687/1; 2734/3; 2799/3; 2841/3; 2891/2; 2932/3;  
 C:Superfamily: dynein heavy chain, cytosolic

Query Match 68.9%; Score 31; DB 2; Length 4131;  
 Best Local Similarity 66.7%; Pred. No. 9.5e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMAQEALAF 9  
 :||:||||  
 Db 598 LMLEEALAF 606

## RESULT 15

AB1360  
 protein gp21 [Bacteriophage A118] [imported] - Listeria monocytogenes (strain EGD-e)  
 C:Species: Listeria monocytogenes  
 C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
 C:Accession: AB1360  
 R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec  
 ; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl,  
 D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001  
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma  
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,  
A:Title: Comparative genomics of *Listeria* species.  
A:Reference number: ABI077; MUID:21537279; PMID:11679669  
A:Accession: ABI360  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-105 <GLA>  
A:Cross-references: GB:NC\_003210; PIDN:CAD00360.1; PID:g16411752; GSPDB:GN00177  
A:Experimental source: strain EGD-e  
C:Genetics:  
A:Gene: lmo2282

Query Match 66.7%; Score 30; DB 2; Length 105;  
Best Local Similarity 70.0%; Pred. No. 34;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 1 LMAQEALAPL 10  
I: ||||| I  
Db 90 LITQELAAAL 99

Search completed: October 7, 2003, 13:32:36  
Job time : 6.63291 secs

**THIS PAGE BLANK (USPTO)**



```
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein HI0636.
GN HI0636.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McClellan K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RA "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512(1995).
CC -1- SIMILARITY: BELONGS TO THE YPCZ/YIIS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U32746; AAC22295.1; -.
CC PIR: I64010; I64010.
CC TIGR: HI0636; -.
CC InterPro: IPR005272; Cons_hypoth743.
CC Pfam: PF04175; DUF406; 1.
CC TIGRFAMs: TIGR00743; TIGR00743; 1.
CC Hypothetical protein; Complete proteome.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 96 AA; 10966 MW; 7719CB9CFD29587F CRC64;
-----
Query Match 71.1%; Score 32; DB 1; Length 96;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 3 AQEALAF 9
DB 42 AQEALAF 48
-----
RESULT 3
PTMA_BACHD
ID PTMA_BACHD STANDARD; PRT; 145 AA.
AC Q9K680;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE PTS system, mannitol-specific IIA component (ETIIA-Mtl) (Mannitol-
DE permease IIA component) (Phosphotransferase enzyme II, A component)
DE (EC 2.7.1.69) (ETII-Mtl).
DE MTLF OR BH3852.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
-----
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -1- FUNCTION: THIS IS A COMPONENT OF THE PHOSPHOENOLPYRUVATE-DEPENDENT
CC SUGAR PHOSPHOTRANSFERASE SYSTEM (PTS), A MAJOR CARBOHYDRATE ACTIVE
CC -TRANSPORT SYSTEM. THE IICD DOMAINS CONTAIN THE SUGAR BINDING SITE
CC AND THE TRANSMEMBRANE CHANNEL; THE IIA DOMAIN CONTAINS THE PRIMARY
CC PHOSPHORYLATION SITE (THE DONOR IS PHOSPHO-HPR); IIA TRANSFERS ITS
CC PHOSPHORYL GROUP TO THE IIB DOMAIN WHICH FINALLY TRANSFERS IT TO
CC THE SUGAR (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Protein N-phosphohistidine + sugar -> protein
CC histidine + sugar phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Contains 1 PTS EIIA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AP001520; BAB07571.1; -.
CC PIR: D84131; D84131.
CC HSPSP; P00550; 1A3A.
CC InterPro: IPR002178; PTS_EIIA_2.
CC Pfam: PF00359; PTS_EIIA_2; 1.
CC PRODOM: PD001689; PTS_EIIA_2; 1.
CC PROSITE: PS00372; PTS_EIIA_2; 1.
CC Phosphotransferase system; Sugar transport; Transferase;
KW Phosphorylation; Complete proteome.
FT MOD_RES 64 64 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 145 AA; 16015 MW; 888B849FA9C9B7FD CRC64;
-----
Query Match 71.1%; Score 32; DB 1; Length 145;
Best Local Similarity 87.5%; Pred. No. 6.6;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 3 AQEALAF 10
DB 133 AQEALAF 140
-----
RESULT 4
VCAC_LAMBD
ID VCAC_LAMBD STANDARD; PRT; 439 AA.
AC P03711;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Minor capsid protein C (GPC) [Contains: Capsid assembly protein NU3].
DE C AND NU3.
OS Bacteriophage lambda.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC Lambda-like viruses.
OX NCBI_TaxID=10710;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83189071; PubMed=6221115;
RA Sanger F., Coulson A.R., Hong G.F., Hill D.F., Petersen G.B.;
RA "Nucleotide sequence of bacteriophage lambda DNA.";
RL J. Mol. Biol. 162:729-773(1982)
CC -1- FUNCTION: IN CONJUNCTION WITH HOST PROTEIN GROE, GENE C PROTEIN
CC SPECIFICALLY CLEAVES SCAFFOLD PROTEIN NU3 DURING ITS REMOVAL
CC FROM THE MATURING PROHEAD.
CC -1- FUNCTION: GENE NU3 PROTEIN FORMS THE SCAFFOLD FOR CAPSID ASSEMBLY
CC AND IS REQUIRED FOR BINDING OF THE B AND C PROTEIN PRODUCTS. IT IS
CC SUBSEQUENTLY LOST FROM THE HEAD DURING MATURATION.
CC -1- MISCELLANEOUS: GENE NU3 IS BELIEVED TO BE CODED WITHIN GENE C, IN
CC THE SAME READING FRAME. THERE ARE TWO POSSIBLE START SITES FOR
CC GENE NU3, THE GTG CODON CORRESPONDING TO 260-VAL AND THE CODON FOR
CC 309-MET.
```



CC -!- MISCELLANEOUS: DURING HEAD ASSEMBLY, THE GENE C PROTEIN IS  
 CC COVALENTLY LINKED WITH AN EQUI-MOLAR AMOUNT OF PROTEIN E AND  
 CC CLEAVED TO YIELD THE MINOR CAPSID PROTEINS X1 AND X2.  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S49.  
 CC -!- CAUTION: THERE ARE FIVE ALTERNATIVE INITIATION SITES FOR GENE C  
 CC TRANSLATION; THE CODON CORRESPONDING TO 1-MET IS THE MOST  
 CC PROBABLE SITE.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; J02459; AAA96537.1; .  
 CC EMBL; J02459; AAA96538.1; ALT\_INIT.  
 CC PIR; A04333; VHBPL.  
 CC MEROPS; S49.003; .  
 CC InterPro; IPR0021142; Peptidase\_U7.  
 CC Pfam; PF01343; Peptidase\_U7; 1.  
 CC ProDom; PD002897; Peptidase\_U7; 1.  
 CC Protease; Hydrolase; Coat protein.  
 CC CHAIN 1 439 MINOR CAPSID PROTEIN C.  
 CC ? 439 CAPSID ASSEMBLY PROTEIN NU3.  
 CC SEQUENCE 439 AA; 45939 MW; 355ED56D60F6CDDF CRC64;  
 CC -----  
 CC Query Match 68.9%; Score 31; DB 1; Length 439;  
 CC Best Local Similarity 80.0%; Pred. No. 35;  
 CC Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 CC -----  
 CC QY 1 LMAQEALAF 10  
 CC | | | | | | |  
 CC Db 52 LTAQEALAF 61  
 CC -----  
 CC RESULT 5  
 CC YKFO\_YEAST  
 CC ID YKFO\_YEAST STANDARD; PRT; 922 AA.  
 CC AC P35736;  
 CC DT 01-JUN-1994 (Rel. 29, Created)  
 CC DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 CC DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 CC DE Hypothetical 103.1 kDa protein in NUP120-CSE4 intergenic region.  
 CC GN YKLO50C OR YKL301 OR YKL263.  
 CC OS Saccharomyces cerevisiae (Baker's yeast).  
 CC OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 CC OX NCBI\_TaxID=4932;  
 CC [1]  
 CC SEQUENCE OF 1-452 FROM N.A.  
 CC RC STRAIN-S288c;  
 CC RX MEDLINE-94205268; PubMed-8154189;  
 CC RA Purnelle B., Skala J., van Dyck L., Tettelin H., Goffeau A.;  
 CC RT "The sequence of a 17.5 kb DNA fragment on the left arm of yeast  
 CC RT chromosome XI identifies the protein kinase gene ELM1, the DNA  
 CC RT primase gene PRI2, a new gene encoding a putative histone and seven  
 CC RT new open reading frames.";  
 CC RL Yeast 9:1379-1384(1993).  
 CC [2]  
 CC SEQUENCE OF 410-922 FROM N.A.  
 CC RC STRAIN-S288c;  
 CC RX MEDLINE-94378723; PubMed-8091862;  
 CC RA Rasmussen S.W.;  
 CC RT "Sequence of a 28.6 kb region of yeast chromosome XI includes the  
 CC RT FBAL and ROA2 genes, an open reading frame (ORF) similar to a  
 CC RT translationally controlled tumour protein, one ORF containing motifs  
 CC RT also found in plant storage proteins and 13 ORFs with weak or no  
 CC RT homology to known proteins.";  
 CC RL Yeast 10:S63-S68(1994).  
 CC -!- SIMILARITY: TO YEAST YNR031C.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X75781; CAA53420.1; .  
 CC EMBL; X71621; .; NOT\_ANNOTATED\_CDS.  
 CC EMBL; Z28050; CAA81855.1; .  
 CC PIR; S37871; S37871.  
 CC SCD; S0001533; YKL050C.  
 CC Hypothetical protein.  
 CC SEQUENCE 922 AA; 103143 MW; 362EE9A3F642DC8B CRC64;  
 CC -----  
 CC Query Match 68.9%; Score 31; DB 1; Length 922;  
 CC Best Local Similarity 77.8%; Pred. No. 77;  
 CC Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 CC -----  
 CC QY 1 LMAQEALAF 9  
 CC | | | | | | |  
 CC Db 173 LAAQKALAF 181  
 CC -----  
 CC RESULT 6  
 CC SMC1\_YEAST  
 CC ID SMC1\_YEAST STANDARD; PRT; 1225 AA.  
 CC AC P32908;  
 CC DT 01-OCT-1993 (Rel. 27, Created)  
 CC DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 CC DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 CC DE Structural maintenance of chromosome 1 (DA-box protein SMC1).  
 CC GN SMC1 OR CHL10 OR YFL008W  
 CC OS Saccharomyces cerevisiae (Baker's yeast).  
 CC OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 CC OX NCBI\_TaxID=4932;  
 CC [1]  
 CC SEQUENCE FROM N.A. AND MUTANTS SMC1-1 AND SMC1-2.  
 CC RX MEDLINE-94103320; PubMed-8276886;  
 CC RA Strunnikov A.V., Larionov V.L., Koshland D.;  
 CC RT "SMC1: an essential yeast gene encoding a putative head-rod-tail  
 CC RT protein is required for nuclear division and defines a new ubiquitous  
 CC RT protein family.";  
 CC RL J. Cell Biol. 123:1635-1648(1993).  
 CC [2]  
 CC SEQUENCE FROM N.A.  
 CC RC STRAIN-S288c / AB972;  
 CC RX MEDLINE-95400292; PubMed-7670463;  
 CC RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,  
 CC RA Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,  
 CC RA Yamazaki M., Tashiro H., Eki T.;  
 CC RT "Analysis of the nucleotide sequence of chromosome VI from  
 CC RT Saccharomyces cerevisiae.";  
 CC RL Nat. Genet. 10:261-268(1995).  
 CC [3]  
 CC IDENTIFICATION IN A COHESIN COMPLEX WITH SMC3; IRR1 AND MCD1, AND  
 CC RP INTERACTION OF THE COHESIN COMPLEX WITH SCC2.  
 CC RX PubMed-9990856;  
 CC RA Toth A., Ciosk R., Uhlmann F., Galova M., Schleiffer A., Nasmyth K.;  
 CC RT "Yeast cohesin complex requires a conserved protein, Eco1p(Ctf7), to  
 CC RT establish cohesion between sister chromatids during DNA replication.";  
 CC RL Genes Dev. 13:320-333(1999).  
 CC [4]  
 CC IDENTIFICATION IN A COHESIN COMPLEX WITH SMC3; MCD1 AND IRR1, AND  
 CC RP STRUCTURE.  
 CC RX PubMed-11983169;  
 CC RA Haering C.H., Loewe J., Hochwagen A., Nasmyth K.;  
 CC RT "Molecular architecture of SMC proteins and the yeast cohesin  
 CC RT complex.";  
 CC RL Mol. Cell 9:773-788(2002).  
 CC -!- FUNCTION: Involved in chromosome cohesion during cell cycle and in

DNA repair. Central component of cohesin complex. The cohesin complex is required for the cohesion of sister chromatids after DNA replication. The cohesin complex apparently forms a large proteinaceous ring within which sister chromatids can be trapped. At anaphase, the complex is cleaved and dissociates from chromatins, allowing sister chromatids to segregate.

-1- SUBUNIT: Cohesin complexes are composed of the SMC1 and SMC3 heterodimer attached to their hinge domain, MCD1/SCC1 which link them, and IRR1/SCC3, which interacts with MCD1. The cohesin complex also interacts with SCC2, which is required for its association with chromosomes.

-1- SUBCELLULAR LOCATION: Nuclear protein. Associates with chromatins. Before prophase it is scattered along chromosome arms. At anaphase, the MCD1 subunit of the cohesin complex is cleaved, leading to the dissociation of the complex from chromosomes, allowing chromosome separation.

-1- DOMAIN: The flexible hinge domain, which separates the large intramolecular coiled coil regions, allows the heterotypic interaction with the corresponding domain of SMC3, forming a V-shaped heterodimer. The two heads of the heterodimer are then connected by different ends of the cleavable MCD1 protein, forming a ring structure.

-1- SIMILARITY: Belongs to the SMC family. SMC1 subfamily.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; L00602; AAA16595.1; -  
EMBL; D50617; BAA09230.1; -  
PIR; A49464; A49464.  
SGD; S0001886; SMC1.  
GO; GO:0008278; C:cohesin complex; IDA.  
GO; GO:0003680; F:AT DNA binding activity; IDA.  
GO; GO:000217; F:DNA secondary structure binding activity; IDA.  
GO; GO:0003690; F:double-stranded DNA binding activity; IDA.  
GO; GO:0000070; P:mitotic chromosome segregation; IMP.  
InterPro; IPR003405; SMC\_C.  
InterPro; IPR003395; SMC\_N.  
Pfam; PF02463; SMC\_N; 1.  
Pfam; PF02483; SMC\_C; 1.  
MitoSys; Cell cycle; Chromosome partition; ATP-binding; Coiled coil; Nuclear protein.  
NP\_BIND 33 40 ATP (POTENTIAL).  
FT DOMAIN 173 489 COILED COIL (POTENTIAL).  
FT DOMAIN 490 678 FLEXIBLE HINGE.  
FT DOMAIN 679 1063 COILED COIL (POTENTIAL).  
FT DOMAIN 1057 1061 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
FT DOMAIN 1129 1164 ALA/ASP-RICH (DA-BOX).  
FT MUTAGEN 173 173 S->L: IN TS MUTANT SMC1-2.  
FT MUTAGEN 458 458 N->D: IN TS MUTANT SMC1-1.  
SQ SEQUENCE 1225 AA; 141279 MW; B504017AAUECCA8C CRC64;

Query Match 68.9%; Score 31; DB 1; Length 1225;  
Best Local Similarity 66.7%; Pred. No. 1e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 MAQEALAF 10  
: : : : :  
Db 566 VAQECIAFL 574

RESULT 7  
HPAI\_SALDU STANDARD; PRT; 263 AA.  
AC Q9RPV0;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)

2,4-dihydroxyhept-2-ene-1,7-dioic acid aldolase (EC 4.1.2.-) (HHED aldolase).  
HPAI.  
ON Salmonella dublin.  
OS Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Salmonella.  
OX NCBI\_TaxID=98360;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=2229;  
RA Galyov E.E., Wood M.W., Hedges S.;  
RT "Characterization of the hpa genetic locus from Salmonella dublin."; Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
RL -1- PATHWAY: 4-hydroxyphenylacetate catabolism; last step.  
CC -1- SIMILARITY: BELONGS TO THE HPCH/HPAI ALDOLASE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AF144422; AAD53496.1; -  
DR HSSP; P23522; 1DXE.  
DR InterPro; IPR005000; Hpch\_HpaI.  
DR Pfam; PF03328; Hpch\_HpaI; 1.  
KW Aromatic hydrocarbons catabolism; Lyase.  
SQ SEQUENCE 263 AA; 27806 MW; 46F34BSC15F6B9A9 CRC64;

Query Match 66.7%; Score 30; DB 1; Length 263;  
Best Local Similarity 87.5%; Pred. No. 35;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMAQEALA 8  
: : : : :  
Db 212 LMANEALA 219

RESULT 8  
NADE\_STRP3 STANDARD; PRT; 274 AA.  
ID NADE\_STRP3  
AC Q8K6D4;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE NH(3)-dependent NAD(+) synthetase (EC 6.3.5.1).  
GN NADE OR SPYM3\_1392 OR SPS0470.  
OS Streptococcus pyogenes (serotype M3).  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=198466;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MGAS315 / Serotype M3;  
FX MEDLINE=22133808; PubMed=12122206;  
RA Beres S.B., Sylva G.L., Barbian K.D., Lei B., Hoff J.S.,  
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Perkins L.D.,  
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,  
RA Schlievert P.M., Musser J.M.;  
RT "Genome sequence of a serotype M3 strain of group A Streptococcus: phage-encoded toxins, the high-virulence phenotype, and clone emergence.";  
RT Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SSI-1 / Serotype M3;  
RA Nakagawa I., Kurokawa K., Nakata M., Tomiyasu Y., Yamashita A.,  
RA Yamazaki K., Okahashi N., Kawabata S., Yasunaga T., Hattori M.,  
RA Hayashi H., Hamada S.;  
RT "The genome of invasive Streptococcus pyogenes: a comparative analysis of S. pyogenes SSI-1, SF370 and MGAS8232.";

```

RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ATP + deamido-NAD(+) + NH(3) + H(2)O = AMP +
CC diphosphate + NAD(+).
CC -!- PATHWAY: NAD biosynthesis.
CC -!- SIMILARITY: BELONGS TO THE NAD SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE014163; AAM79999.1; -.
DR EMBL; AP005142; BAC63565.1; -.
DR HAMAP; MF_00193; -.
DR InterPro; IPR003694; NAD_synthase.
DR Pfam; PF02540; NAD_synthase; 1.
DR TIGRFAMs; TIGR00552; nade; 1.
DR KW Ligase; NAD; ATP-binding; Complete proteome.
DR NP_BIND 46 53 ATP (BY SIMILARITY).
DR FT ACT_SITE 48 48 BY SIMILARITY.
DR SQ SEQUENCE 274 AA; 30180 MW; C52B01B7F9D40B87 CRC64;

Query Match 66.7%; Score 30; DB 1; Length 274;
Best Local Similarity 75.0%; Pred. No. 36;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 AQEALAFI 10
DB 94 AQKALAFI 101

RESULT 9
NADE_STRPN STANDARD; PRT; 274 AA.
AC Q8NY8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE NH(3)-dependent NAD(+) synthetase (EC 6.3.5.1).
GN NADE OR SPYM18.1663.
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OC NCBI_TaxID=186103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108;
RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RA "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
CC -!- CATALYTIC ACTIVITY: ATP + deamido-NAD(+) + NH(3) + H(2)O = AMP +
CC diphosphate + NAD(+).
CC -!- PATHWAY: NAD biosynthesis.
CC -!- SIMILARITY: BELONGS TO THE NAD SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE010078; AAL98207.1; -.

```

```

DR HAMAP; MF_00193; -.
DR InterPro; IPR003694; NAD_synthase.
DR Pfam; PF02540; NAD_synthase; 1.
DR TIGRFAMs; TIGR00552; nade; 1.
DR KW Ligase; NAD; ATP-binding; Complete proteome.
DR NP_BIND 46 53 ATP (BY SIMILARITY).
DR FT ACT_SITE 48 48 BY SIMILARITY.
DR SQ SEQUENCE 274 AA; 30208 MW; DA7CD7F2004BB3C7 CRC64;

Query Match 66.7%; Score 30; DB 1; Length 274;
Best Local Similarity 75.0%; Pred. No. 36;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 AQEALAFI 10
DB 94 AQKALAFI 101

RESULT 10
NADE_STRPN STANDARD; PRT; 274 AA.
AC Q97QI3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE NH(3)-dependent NAD(+) synthetase (EC 6.3.5.1).
GN NADE OR SPI420 OR SPRI276.
OS Streptococcus pneumoniae, and
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OC NCBI_TaxID=1313, 171101;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-334 / TIGR4;
RX MEDLINE=21357209; PubMed=11463916;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Maynam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae";
RL Science 293:498-506(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-255 / R6;
RX MEDLINE=21429245; PubMed=11544234;
RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,
RA DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
RA Gilmour R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
RA McAhren S.M., McHenney M., McLeaster K., Mundy C.W., Nicas T.I.,
RA Norris F.H., O'Garra M., Peery R.B., Robertson G.T., Rocky P.,
RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostock P.R. Jr., Skatrud P.L.,
RA Glass J.I.;
RT "Genome of the bacterium Streptococcus pneumoniae strain R6.";
RL J. Bacteriol. 183:5709-5717(2001).
CC -!- CATALYTIC ACTIVITY: ATP + deamido-NAD(+) + NH(3) + H(2)O = AMP +
CC diphosphate + NAD(+).
CC -!- PATHWAY: NAD biosynthesis.
CC -!- SIMILARITY: BELONGS TO THE NAD SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE007439; AAK75517.1; -.
DR EMBL; AE008499; AAL00080.1; -.
DR PIR; C98031; C98031.
DR PIR; D95165; D95165.
DR TIGR; SP1420; -.
DR HAMAP; MF_00193; -.
DR InterPro; IPR003694; NAD_synthase.
DR Pfam; PF02540; NAD_synthase; 1.
DR TIGRFAMs; TIGR00552; nade; 1.
KW Ligase; NAD; ATP-binding; Complete proteome.
FT NP_BIND 46 53 ATP (BY SIMILARITY).
FT ACT_SITE 48 48 BY SIMILARITY.
SQ SEQUENCE 274 AA; 30135 MW; 10972D8AF4A9044C CRC64;

Query Match 66.7%; Score 30; DB 1; Length 274;
Best Local Similarity 75.0%; Pred. No. 36;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 AQEALAFI 10
DB 94 AOKALAFI 101

RESULT 11
ID NAME_STRPY STANDARD; PRT; 274 AA.
AC Q99YK9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE NH(3)-dependent NAD(+) synthetase (EC 6.3.5.1).
GN NADE OR SPY1652.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
ON NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
CC -|- CATALYTIC ACTIVITY: ATP + deamido-NAD(+) + NH(3) + H(2)O = AMP +
CC diphosphate + NAD(+).
CC -|- PATHWAY: NAD biosynthesis.
CC -|- SIMILARITY: BELONGS TO THE NAD SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE006596; AAK34418.1; -.
DR HSSP; P08164; 2NSY.
DR HAMAP; MF_00193; -.
DR InterPro; IPR003694; NAD_synthase.
DR Pfam; PF02540; NAD_synthase; 1.
DR TIGRFAMs; TIGR00552; nade; 1.
KW Ligase; NAD; ATP-binding; Complete proteome.
FT NP_BIND 46 53 ATP (BY SIMILARITY).
FT ACT_SITE 48 48 BY SIMILARITY.
SQ SEQUENCE 274 AA; 30074 MW; 79B336998F76B0DC CRC64;

Query Match 66.7%; Score 30; DB 1; Length 274;

Best Local Similarity 75.0%; Pred. No. 36;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 AQEALAFI 10
DB 94 AOKALAFI 101

RESULT 12
BLOZ_SALTY STANDARD; PRT; 275 AA.
AC P05191; Q57015;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE beta-lactamase OXA-2 precursor (EC 3.5.2.6) (Penicillinase).
GN OXA2.
OS Salmonella typhimurium, and
OS Escherichia coli.
OG Plasmid IncN R46, and Plasmid pBP11.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OC NCBI_TaxID=602, 562;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; STRAIN=Type 1A; PLASMID=IncN R46;
RX MEDLINE=86030657; PubMed=3876949;
RA Dale J.W., Godwin D., Mossakowska D., Stephenson P., Wall S.;
RT "Sequence of the OXA2 beta-lactamase: comparison with other
RT penicillin-reactive enzymes.";
RL FEBS Lett. 191:39-44(1985).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; PLASMID=IncN R46;
RX MEDLINE=89170730; PubMed=2538329;
RA Mossakowska D., Ali N.A., Dale J.W.;
RT "Oxacillin-hydrolysing beta-lactamases. A comparative analysis at
RT nucleotide and amino acid sequence levels.";
RL Eur. J. Biochem. 180:309-318(1989).
RN [3]
RP SEQUENCE OF 22-32.
RC SPECIES=S.typhimurium; PLASMID=IncN R46;
RX MEDLINE=85121787; PubMed=6335398;
RA Holland S., Dale J.W.;
RT "Improved purification and characterization of the OXA-2
RT beta-lactamase.";
RL Biochem. J. 224:1009-1013(1984).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; PLASMID=OBP11;
RX MEDLINE=90095432; PubMed=2689593;
RA Nuecken E.J., Henschke R.B., Schmidt F.R.J.;
RT "Nucleotide sequence of an OXA-2 beta-lactamase gene from the
RT R-plasmid R1767 derived plasmid pBP11 and comparison to closely
RT related resistance determinants found in R46 and Tn2603.";
RL J. Gen. Microbiol. 135:761-765(1989).
RN [5]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; PLASMID=IncN R46;
RX MEDLINE=88015610; PubMed=2821509;
RA Hall R.M., Vockler C.;
RT "The region of the IncN plasmid R46 coding for resistance to
RT beta-lactam antibiotics, streptomycin/spectinomycin and sulphonamides
RT is closely related to antibiotic resistance segments found in IncW
RT plasmids and in Tn21-like transposons.";
RL Nucleic Acids Res. 15:7491-7501(1987).
RN [6]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; PLASMID=IncN R46;
RX MEDLINE=93096861; PubMed=1334268;
RA Stokes H.W., Hall R.M.;
RT "The integron InI in plasmid R46 includes two copies of the oxa2 gene
RT cassette.";
```

```
RL Plasmid 28:225-234(1992).
CC -!- FUNCTION: THIS IS AN OXACILLIN HYDROLYZING BETA-LACTAMASE.
CC -!- CATALYTIC ACTIVITY: A beta-lactam + H(2)O = a substituted beta-
CC amino acid.
CC
CC -!- SIMILARITY: BELONGS TO THE CLASS-D BETA-LACTAMASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X03037; CAA26839.1; -
DR EMBL; X07260; CAA30246.1; -
DR EMBL; M25261; AAB98357.1; -
DR EMBL; M95287; AAB59082.1; -
DR EMBL; M95287; AAB59084.2; -
DR PIR; A91350; PNEBT.
DR HSSP; P14489; 1E3U.
DR InterPro; IPR002137; Beta_lactamase_D.
DR PROSITE; PS00337; BETA_LACTAMASE_D; 1.
KW Hydrolase; Antibiotic resistance; Signal; Plasmid.
FT SIGNAL 1 21
FT CHAIN 22 275
FT ACT_SITE 72 72
FT BINDING 210 212
FT SEQUENCE 275 AA; 31686 MW; D3678DFC878DE8B6 CRC64;
SQ
Query Match 66.7%; Score 30; DB 1; Length 275;
Best Local Similarity 60.0%; Pred. No. 36;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
OY 1 LMAQEQALAF 10
DB 166 ISAQEQIAFL 175
: ||| :|||
RESULT 13
BLO3_PSEAE STANDARD; PRT; 275 AA.
ID BLO3_PSEAE
AC Q51429;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Beta-lactamase OXA-3 precursor (EC 3.5.2.6) (Penicillinase).
GN OXA3.
OS Pseudomonas aeruginosa.
OG Plasmid pMG25.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95305545; PubMed=7785990;
RA Sanschagrin F., Couture F., Levesque R.C.;
RT "Primary structure of OXA-3 and phylogeny of oxacillin-hydrolyzing
RT class D beta-lactamases.";
RL Antimicrob. Agents Chemother. 39:887-893(1995).
CC -!- FUNCTION: THIS IS AN OXACILLIN-HYDROLYZING BETA-LACTAMASE.
CC -!- CATALYTIC ACTIVITY: A beta-lactam + H(2)O = a substituted beta-
CC amino acid.
CC -!- SIMILARITY: BELONGS TO THE CLASS-D BETA-LACTAMASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

```
DR EMBL; L07945; AAC411449.1; -
DR HSSP; P14489; 1E3U.
DR InterPro; IPR002137; Beta_lactamase_D.
DR PROSITE; PS00337; BETA_LACTAMASE_D; 1.
KW Hydrolase; Antibiotic resistance; Signal; Plasmid.
FT SIGNAL 1 21
FT CHAIN 22 275
FT ACT_SITE 72 72
FT BINDING 210 212
FT SEQUENCE 275 AA; 31879 MW; 4F9F95D755140146 CRC64;
SQ
Query Match 66.7%; Score 30; DB 1; Length 275;
Best Local Similarity 60.0%; Pred. No. 36;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
OY 1 LMAQEQALAF 10
DB 166 TAAQEQIAFL 175
: ||| :|||
RESULT 14
BLOF_PSEAE STANDARD; PRT; 275 AA.
ID BLOF_PSEAE
AC Q51574;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Beta-lactamase OXA-15 precursor (EC 3.5.2.6).
GN OXA15.
OS Pseudomonas aeruginosa.
OG Plasmid pMLH54.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX STRAIN-AH;
RX MEDLINE=97242491; PubMed=9087490;
RA Danel F., Hall L.M.C., Gur D., Livermore D.M.;
RT "OXA-15, an extended-spectrum variant of OXA-2 beta-lactamase,
RT isolated from a Pseudomonas aeruginosa strain.";
RL Antimicrob. Agents Chemother. 41:785-790(1997).
CC -!- FUNCTION: HYDROLYZES OXACILLIN, FIRST-GENERATION CEPHALOSPORINS
CC AND CEFOTAZIDIME. DOES NOT HYDROLYZE CEFOTAXIME OR CARBAPENEMS.
CC -!- CATALYTIC ACTIVITY: A beta-lactam + H(2)O = a substituted beta-
CC amino acid.
CC -!- SIMILARITY: BELONGS TO THE CLASS-D BETA-LACTAMASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U63835; AAB05874.1; -
DR HSSP; P14489; 1E3U.
DR InterPro; IPR002137; Beta_lactamase_D.
DR PROSITE; PS00337; BETA_LACTAMASE_D; 1.
KW Hydrolase; Antibiotic resistance; Signal; Plasmid.
FT SIGNAL 1 21
FT CHAIN 22 275
FT ACT_SITE 72 72
FT BINDING 210 212
FT SEQUENCE 275 AA; 31628 MW; D3678DFE678C38B5 CRC64;
SQ
Query Match 66.7%; Score 30; DB 1; Length 275;
Best Local Similarity 60.0%; Pred. No. 36;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
OY 1 LMAQEQALAF 10
DB 166 TAAQEQIAFL 175
: ||| :|||
```

Tue Oct 7 15:52:25 2003

Db 166 ISAQEQIAFL 175

## RESULT 15

DPO4\_LACLA STANDARD; PRT; 363 AA.

AC Q9CE21;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE DNA polymerase IV (EC 2.7.7.7) (Pol IV).  
GN DINB OR LL2029.  
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.  
OX NCBI\_TaxID=1360;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=IL1403;  
RX MEDLINE=21235186; PubMed=11337471;  
RA Bolotin A., Wincker P., Mauer S., Jaillon O., Malarne K.,  
RA Weissenbach J., Ehrlich S.D., Sorokin A.;  
RA "The complete genome sequence of the lactic acid bacterium Lactococcus  
RT lactis ssp. lactis IL1403.";  
RL Genome Res. 11:731-753(2001).  
CC -1- FUNCTION: Poorly processive, error-prone DNA polymerase involved  
CC in untargeted mutagenesis. Copies undamaged DNA at stalled  
CC replication forks, which arise in vivo from mismatched or  
CC misaligned primer ends. These misaligned primers can be extended  
CC by polIV. Exhibits no 3'-5' exonuclease (proofreading) activity.  
CC May be involved in translesional synthesis, in conjunction with  
CC the beta clamp from polIII (By similarity).  
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate  
CC + [DNA](N).  
CC -1- COFACTOR: Binds two magnesium ions (By similarity).  
CC -1- SUBUNIT: Monomer (By similarity).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-Y FAMILY.  
CC -1- SIMILARITY: Contains 1 umuC domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AE006432; AAK06127.1; -.  
CC PIR; E86878; E86878.  
CC HAMAP; MF\_01113; -; 1.  
CC InterPro; IPR001126; UMUC\_like.  
CC Pfam; PF00817; IMS; 1.  
CC PROSITE; PS50173; UMUC; 1.  
CC Transferrase; DNA-directed DNA polymerase; DNA replication; DNA repair;  
CC DNA-binding; Mutator protein; Magnesium; Complete proteome.  
CC DOMAIN 14 197  
CC ACT SITE 23 23 SUBSTRATE DISCRIMINATION (BY SIMILARITY).  
CC ACT SITE 116 116 BY SIMILARITY.  
CC METAL 18 18 MAGNESIUM (BY SIMILARITY).  
CC METAL 115 115 MAGNESIUM (BY SIMILARITY).  
CC SEQUENCE 363 AA; 40x38 MW; 4E4DFA9C0CEC168B CRC64;  
Query Match 66.7%; Score 30; DB 1; Length 363;  
Best Local Similarity 60.0%; Pred. No. 48;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 1 LMAQEAFL 10  
: :| | | |  
Db 176 IMPEALEFL 185

Search completed: October 7, 2003, 13:26:24  
Job time : 3.91139 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:18:33 ; Search time 16.7089 Seconds  
(without alignments)  
154.441 Million cell updates/sec

Title: US-09-807-512-12  
Perfect score: 45  
Sequence: 1 LMAQEAFL 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	58	4 O95146	095146 homo sapien
2	45	100.0	109	4 O95987	095987 homo sapien
3	34	75.6	654	16 Q8DGN6	Q8dgn6 synecococc
4	33	73.3	318	10 Q8SBE4	Q8sbe4 arabidopsis
5	33	73.3	320	10 Q8LFR4	Q8lfr4 arabidopsis
6	33	73.3	333	10 Q9SJ50	Q9sj50 arabidopsis
7	33	73.3	338	4 Q8WU54	Q8wu54 homo sapien
8	33	73.3	782	4 Q86A82	Q86a82 homo sapien
9	32	71.1	205	16 Q9KVK7	Q9kvk7 vibrio chol
10	32	71.1	215	10 Q9LXH8	Q9lyh8 arabidopsis
11	32	71.1	373	16 Q8FP51	Q8fp51 corynebacte
12	32	71.1	592	2 Q9S433	Q9s433 myxococcu
13	32	71.1	602	16 Q9Q013	Q9qq13 streptomyce
14	32	71.1	2692	5 Q95Q95	Q95q95 caenorhabd
15	32	71.1	2695	5 O01438	O01438 caenorhabd
16	31	68.9	143	10 Q9XH06	Q9xh06 arabidopsis

17	31	68.9	350	16 Q92RG8	Q92rg8 rhizobium m
18	31	68.9	398	17 Q30103	Q30103 archaeoglob
19	31	68.9	399	10 Q9ZWJ5	Q9zwj5 zea mays (m
20	31	68.9	435	16 Q98EL2	Q98el2 rhizobium l
21	31	68.9	477	16 Q8FEV8	Q8fev8 escherichia
22	31	68.9	636	16 Q9JSJ0	Q9jsj0 chlamydia p
23	31	68.9	636	16 Q92971	Q92971 chlamydia p
24	31	68.9	1316	3 Q8TFG6	Q8tf6g schizosacch
25	31	68.9	2538	2 Q93RN7	Q93rn7 xenorhabdus
26	31	68.9	4171	5 Q19542	Q19542 caenorhabdi
27	31	68.9	6266	2 Q8GM85	Q8gm85 symbiont ba
28	30	66.7	105	9 Q9TIA2	Q9tia2 bacterioph
29	30	66.7	105	16 Q8Y4Z7	Q8y4z7 listeria mo
30	30	66.7	108	16 Q9CLX9	Q9clx9 pasteurella
31	30	66.7	196	4 Q9NQF0	Q9nqf0 homo sapien
32	30	66.7	238	16 Q8Z3A9	Q8z3a9 salmonella
33	30	66.7	238	16 Q9L6P4	Q9l6p4 s putative
34	30	66.7	240	2 Q47Z39	Q47z39 escherichia
35	30	66.7	246	2 Q9F613	Q9f613 pseudomonas
36	30	66.7	257	2 Q9AFD6	Q9afd6 pseudomonas
37	30	66.7	257	4 Q9HC13	Q9hc13 homo sapien
38	30	66.7	263	16 Q8ZQ47	Q8zq47 salmonella
39	30	66.7	263	16 Q8Z7P9	Q8z7p9 salmonella
40	30	66.7	275	2 P94124	P94124 acinetobact
41	30	66.7	275	2 Q93PR0	Q93pr0 pseudomonas
42	30	66.7	308	16 Q916Z4	Q916z4 pseudomonas
43	30	66.7	317	11 Q8VFH0	Q8vfho mus musculus
44	30	66.7	397	16 Q93J54	Q93j54 streptomyce
45	30	66.7	424	16 Q9KARI	Q9karl bacillus ha

#### ALIGNMENTS

#### RESULT 1

O95146 ID O95146 PRELIMINARY; PRT; 58 AA.  
AC O95146;  
DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE LAGE-2ALT protein isoform.  
GN LAGE-2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Wang R.-F., Johnston S.L., Zeng G., Topalian S.L.,  
RA Schwartzentruber D.J., Rosenberg S.A.;  
RT "A Breast and Melanoma-Shared Tumor Antigen: T Cell Responses to  
RT Antigenic Peptides Translated from Different Open Reading Frames.";  
RL J. Immunol. 161:3596-3606(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Lethe B.G.;  
RN [3]  
RP Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
MEDLINE=99454989; PubMed=10523621;  
RA De Smet C., Lurquin C., Lethe B., Martelange V., Boon T.;  
RT "DNA methylation is the primary silencing mechanism for a set of germ  
RT line- and tumor-specific genes with a CpG-rich promoter.";  
RL Mol. Cell. Biol. 19:7327-7335(1999).  
DR EMBL; AF038567; AAD05203.1; -;  
DR EMBL; AJ275977; CAB76944.1; -;  
SQ SEQUENCE 58 AA; 6188 MW; ED12057564BC7EF2 CRC64;

Query Match 100.0%; Score 45; DB 4; Length 58;  
Best Local Similarity 100.0%; Pred. No. 0.033;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 LMAQEALAF 10
        | : | | | : |
Db     598 LVAOETLAY 607
```



```
Best Local Similarity 77.8%; Pred. NO. 77;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 MAQEALAF 10
DB 1 MQEALSFL 9

RESULT 6
Q9SJ50
ID Q9SJ50 PRELIMINARY; PRT; 333 AA.
AC Q9SJ50;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE At2g36000/F11f19.9.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Gibson H.A., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Yamada K., Yu G., Yu S., Shinozaki K.,
RA Davis R.W., Theologis A., Ecker J.R.;
RA "Arabidopsis cdNA clones.";
RA Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Meyers M.C., Banh J., Bowser L.,
RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RA "Arabidopsis ORF clones.";
RA Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AY039888; AAK63992.1; -
DR EMBL; AY039871; AAK63975.1; -
DR EMBL; AY077666; AAL76144.1; -
DR InterPro; IPR003690; mTERF.
DR Pfam; PF02536; mTERF; 1.
DR SMART; SM00733; mTerf; 5.
SQ SEQUENCE 333 AA; 37952 MW; 072192A86556889A CRC64;

Query Match 73.3%; Score 33; DB 10; Length 333;
Best Local Similarity 77.8%; Pred. NO. 81;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 MAQEALAF 10
DB 1 MQEALSFL 9

RESULT 7
Q8WU54
ID Q8WU54 PRELIMINARY; PRT; 338 AA.
AC Q8WU54;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=El Tor N16961 / Serotype O1;
```

```
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC021238; AAH21238.1; -
DR Pfam; PF04326; DUF467; 1.
KW Hypothetical protein.
SQ SEQUENCE 338 AA; 38102 MW; F321EDDA06C592FE CRC64;

Query Match 73.3%; Score 33; DB 4; Length 338;
Best Local Similarity 87.5%; Pred. NO. 82;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 AQEALAF 10
DB 130 SQEALAF 137

RESULT 8
Q96A82
ID Q96A82 PRELIMINARY; PRT; 782 AA.
AC Q96A82;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ30106 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Nishi T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Ota T., Wakamatsu A.,
RA Ishii S., Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Yamashita H., Matsuo K., Nakamura Y., Sekine M.,
RA Kikuchi H., Kanda K., Wagatsuma M., Murakawa K., Kanehori K.,
RA Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y.,
RA Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RA "NEDO human cDNA sequencing project.";
RA Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AK054668; BAB70788.1; -
DR Pfam; PF04326; DUF467; 1.
KW Hypothetical protein.
FT NON_TER 782
SQ SEQUENCE 782 AA; 89004 MW; 7BA657A17714028B CRC64;

Query Match 73.3%; Score 33; DB 4; Length 782;
Best Local Similarity 87.5%; Pred. NO. 1.9e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 AQEALAF 10
DB 130 SQEALAF 137

RESULT 9
Q9KVK7
ID Q9KVK7 PRELIMINARY; PRT; 205 AA.
AC Q9KVK7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein VC0136.
GN VC0136.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=El Tor N16961 / Serotype O1;
```

RX MEDLINE=20406833; PubMed=10952301;  
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,  
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,  
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,  
 RA McDonald L., Uterback T., Fleischmann R.D., Nierman W.C., White O.,  
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
 RA Fraser C.M.;  
 RA "DNA sequence of both chromosomes of the cholera pathogen *Vibrio*  
 RT *cholerae*.";  
 RL Nature 406:477-483(2000).  
 DR EMBL: AE004104; AAF93313.1; -.  
 DR TIGR: VC0136; -.  
 DR InterPro: IPR001123; Lyse.  
 DR Pfam: PF01810; Lyse; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 205 AA; 22099 MW; AE42CA6FA6EA1223 CRC64;

Query Match 71.1%; Score 32; DB 16; Length 205;  
 Best Local Similarity 77.8%; Pred. No. 81;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMAQALAF 9  
 Db 64 LVAQALAF 72  
 |||||

RESULT 10  
 ID Q9LYH8 PRELIMINARY; PRT; 215 AA.  
 AC Q9LYH8;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Lysine decarboxylase-like protein.  
 GN Flaf18.120.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bevan M., Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A.,  
 RA Bancroft I., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.;  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL163812; CAB87668.1; -.  
 DR InterPro: IPR005269; Cons\_hypoth730.  
 DR Pfam: PF03641; Lysine\_decarbox; 1.  
 DR TIGRFAMs: TIGR00730; TIGR00730; 1.  
 SQ SEQUENCE 215 AA; 23774 MW; A8FFD98BA74F853 CRC64;

Query Match 71.1%; Score 32; DB 10; Length 215;  
 Best Local Similarity 77.8%; Pred. No. 85;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 MAQALAF 10  
 Db 102 MAQALAF 110  
 |||||

RESULT 11  
 ID Q8FP51 PRELIMINARY; PRT; 373 AA.  
 AC Q8FP51;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Putative molybdopterin biosynthesis protein Moeb.

GN CE1937.  
 OS Corynebacterium efficiens.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
 OX NCBI\_TaxID=152794;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;  
 RA Kawarabayasi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,  
 RA Ikeo K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.,  
 RA Usuda Y., Sugimoto S.;  
 RA "The entire genomic sequence of *Corynebacterium efficiens* YS-314.";  
 RT Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AP005220; BAC18747.1; -.  
 KW Complete proteome.  
 SQ SEQUENCE 373 AA; 39668 MW; 08ED72297A8341B3 CRC64;

Query Match 71.1%; Score 32; DB 16; Length 373;  
 Best Local Similarity 70.0%; Pred. No. 1.5e+02;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LMAQALAF 10  
 Db 211 LMATEAKFL 220  
 |||||

RESULT 12  
 ID Q9S433 PRELIMINARY; PRT; 592 AA.  
 AC Q9S433;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Putative membrane protein.  
 GN ESPB.  
 OS *Myxococcus xanthus*.  
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;  
 OC Cystobacterineae; Myxococcaceae; Myxococcus.  
 OX NCBI\_TaxID=34;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DZ2;  
 RX MEDLINE=20032045; PubMed=10564511;  
 RA Cho K., Zusman D.R.;  
 RT "Sporulation timing in *Myxococcus xanthus* is controlled by the *espAB*  
 RT locus.";  
 RL Mol. Microbiol. 34:714-725(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DZ2;  
 RA Cho K., Zusman D.R.;  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF163841; AAD47813.2; -.  
 DR InterPro: IPR004648; Oligopept\_transpt.  
 DR TIGRFAMs: TIGR00728; OPT\_sfam; 1.  
 SQ SEQUENCE 592 AA; 58995 MW; 6E6CC164A6B1D205 CRC64;

Query Match 71.1%; Score 32; DB 2; Length 592;  
 Best Local Similarity 77.8%; Pred. No. 2.4e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMAQALAF 9  
 Db 154 LVAQALPF 162  
 |:|||||

RESULT 13  
 ID Q99Q13 PRELIMINARY; PRT; 602 AA.  
 AC Q99Q13;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)

DE Putative ATP/GTP-binding protein.  
GN SCPI.290C AND SCPI.63.  
OS Streptomyces coelicolor.  
OG Plasmid SCPI.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1902;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RX MEDLINE=21996410; PubMed=12000953;  
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,  
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,  
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
RA Huang C.-H., Kieser T., Larke L., Larke L., Murphy L., O'Neill S.,  
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,  
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,  
RA Hopwood D.A.;  
RT "Complete genome sequence of the model actinomycete Streptomyces  
RT coelicolor A3(2).";  
RL Nature 417:141-147(2002).  
DR EMBL: AL590464; CAC36816.1; -;  
DR EMBL: AL590463; CAC36585.1; -;  
SQ Plasmid; Complete proteome.  
SQ SEQUENCE 602 AA; 65653 MW; D85E86E6C3B216A0 CRC64;

Query Match 71.1%; Score 32; DB 16; Length 602;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 QEALAF 10  
Db 530 QEALAF 536  
|||||

RESULT 14  
Q95Q95 PRELIMINARY; PRT; 2692 AA.  
AC Q95Q95;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-WAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Hypothetical protein B0261.2b.  
GN B0261.2.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RX MEDLINE=99069613; PubMed=9851916;  
RA Waterston R.;  
RT "Genome sequence of the nematode C. elegans: a platform for  
RT investigating biology. The C. elegans Sequencing Consortium.";  
RL Science 282:2012-2018(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Scheet P., Gattung S.;  
RT "The sequence of C. elegans cosmid B0261.";  
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Waterston R.;  
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: U97016; AAN84886.1; -;  
KW Hypothetical protein.  
SQ SEQUENCE 2692 AA; 305847 MW; 80B31C5469D59261 CRC64;

Query Match 71.1%; Score 32; DB 5; Length 2692;  
Best Local Similarity 77.8%; Pred. No. 1.1e+03;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 MAQEALAF 10  
Db 163 MASKALAF 171  
|||||

RESULT 15  
O01438 PRELIMINARY; PRT; 2695 AA.  
AC O01438;  
DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-WAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Hypothetical protein B0261.2a.  
GN B0261.2.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RX MEDLINE=99069613; PubMed=9851916;  
RA Waterston R.;  
RT "Genome sequence of the nematode C. elegans: a platform for  
RT investigating biology. The C. elegans Sequencing Consortium.";  
RL Science 282:2012-2018(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Scheet P., Gattung S.;  
RT "The sequence of C. elegans cosmid B0261.";  
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Waterston R.;  
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: U97016; AAN84885.1; -;  
KW Hypothetical protein.  
SQ SEQUENCE 2695 AA; 306173 MW; D15ADFA7B33825BA CRC64;

Query Match 71.1%; Score 32; DB 5; Length 2695;  
Best Local Similarity 77.8%; Pred. No. 1.1e+03;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 MAQEALAF 10  
Db 163 MASKALAF 171  
|||||

Search completed: October 7, 2003, 13:30:59  
Job time : 17.7089 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:21:07 ; Search time 5.44304 Seconds  
(without alignments)  
77.734 Million cell updates/sec

Title: US-09-807-512-12  
Perfect score: 45  
Sequence: 1 LMAQALAF 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents\_AA.\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pap.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pap.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pap.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pap.\*
- 5: /cgn2\_6/ptodata/1/iaa/PT05\_COMB.pap.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32	71.1	360	US-09-328-352-7879	Sequence 7879, Ap
2	32	71.1	552	US-09-773-426A-9	Sequence 9, Appli
3	31	68.9	268	US-09-252-991A-23138	Sequence 23138, A
4	31	68.9	268	US-09-252-991A-23139	Sequence 23139, A
5	31	68.9	270	US-09-252-991A-23186	Sequence 23186, A
6	31	68.9	503	US-09-252-991A-26258	Sequence 26258, A
7	31	68.9	636	US-09-198-452A-128	Sequence 128, App
8	30	66.7	92	US-08-903-223-408	Sequence 408, App
9	30	66.7	274	US-09-250-677-2	Sequence 2, Appli
10	30	66.7	338	US-09-252-991A-28498	Sequence 28498, A
11	30	66.7	415	US-09-176-657-8	Sequence 8, Appli
12	30	66.7	415	US-09-421-299-8	Sequence 8, Appli
13	30	66.7	463	US-09-082-310-1	Sequence 1, Appli
14	30	66.7	463	US-09-575-205-1	Sequence 1, Appli
15	30	66.7	763	US-09-252-991A-30146	Sequence 30146, A
16	29	64.4	118	US-08-588-258B-38	Sequence 38, Appli
17	29	64.4	118	PCT-US96-08295-38	Sequence 38, Appli
18	29	64.4	119	US-09-328-352-8051	Sequence 8051, Ap
19	29	64.4	153	US-08-936-165A-507	Sequence 507, App
20	29	64.4	225	US-09-198-452A-648	Sequence 648, App
21	29	64.4	307	US-09-252-991A-20828	Sequence 20828, A
22	29	64.4	394	US-08-791-115B-27	Sequence 27, Appli
23	29	64.4	394	US-09-140-749-17	Sequence 17, Appli
24	29	64.4	403	US-08-791-115B-1	Sequence 1, Appli
25	29	64.4	403	US-08-791-115B-23	Sequence 23, Appli
26	29	64.4	403	US-08-791-115B-25	Sequence 25, Appli
27	29	64.4	403	US-09-140-749-2	Sequence 2, Appli

28	29	64.4	403	4	US-09-140-749-10	Sequence 10, Appl
29	29	64.4	403	4	US-09-140-749-49	Sequence 49, Appl
30	29	64.4	403	4	US-09-140-749-55	Sequence 55, Appl
31	29	64.4	403	4	US-09-140-749-57	Sequence 57, Appl
32	29	64.4	430	3	US-08-791-115B-7	Sequence 7, Appli
33	29	64.4	430	4	US-09-140-749-18	Sequence 18, Appl
34	29	64.4	542	4	US-08-311-731A-213	Sequence 213, App
35	29	64.4	559	4	US-09-140-749-15	Sequence 15, Appl
36	29	64.4	565	3	US-08-906-156A-12	Sequence 12, Appli
37	29	64.4	597	4	US-09-140-749-8	Sequence 8, Appli
38	29	64.4	645	3	US-08-791-115B-6	Sequence 6, Appli
39	29	64.4	670	4	US-09-252-991A-24930	Sequence 24930, A
40	29	64.4	729	4	US-09-107-532A-6946	Sequence 6946, Ap
41	29	64.4	742	3	US-08-791-115B-5	Sequence 5, Appli
42	29	64.4	797	1	US-08-698-551-18	Sequence 18, Appl
43	29	64.4	797	2	US-08-839-032A-18	Sequence 18, Appl
44	29	64.4	797	4	US-09-185-258C-18	Sequence 18, Appl
45	29	64.4	1190	1	US-08-337-690A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-328-352-7879  
; Sequence 7879, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 7879  
; LENGTH: 360  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-7879

Query Match 71.1%; Score 32; DB 4; Length 360;  
Best Local Similarity 60.0%; Pred. No. 76;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY	1	LMAQALAF 10
DB	14	IMAQRLSFL 23

RESULT 2

US-09-773-426A-9  
; Sequence 9, Application US/09773426A  
; Patent No. 6534302  
; GENERAL INFORMATION:  
; APPLICANT: Glucksmann, Maria Alexandra  
; APPLICANT: Williamson, Mark  
; APPLICANT: Tsia, Fong-Ying  
; APPLICANT: Rudolph-Owen, Laura A.  
; TITLE OF INVENTION: 22438, 23553, 25278, and 26212 NO. 6534302a1  
; FILE REFERENCE: Human Sulfatases (A CIP Application)  
; CURRENT APPLICATION NUMBER: US/09/773,426A  
; CURRENT FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: US 09/495,823  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9  
; LENGTH: 552  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:

```
; OTHER INFORMATION: Pfam consensus sequence for human sulfatase
US-09-773-426A-9

Query Match          71.1%; Score 32; DB 4; Length 552;
Best Local Similarity 70.0%; Pred. No. 1.2e+02;
Matches              7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 LMAQEALAF 10
      |:| |||||
Db      262 LLAEALPFL 271

RESULT 3
US-09-252-991A-23138
; Sequence 23138, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23138
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23138

Query Match          68.9%; Score 31; DB 4; Length 268;
Best Local Similarity 66.7%; Pred. No. 88;
Matches              6; Conservative 2; Mismatches 1; Indels 1; Gaps 0;

QY      1 LMAQEALAF 9
      |:| |||||
Db      88 MLAEALIF 96

RESULT 4
US-09-252-991A-23139
; Sequence 23139, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23139
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23139

Query Match          68.9%; Score 31; DB 4; Length 268;
Best Local Similarity 66.7%; Pred. No. 88;
Matches              6; Conservative 2; Mismatches 1; Indels 1; Gaps 0;

QY      1 LMAQEALAF 9
      |:| |||||
Db      88 MLAEALIF 96

RESULT 5
US-09-252-991A-23186
; Sequence 23186, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23186
; LENGTH: 270
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23186

Query Match          68.9%; Score 31; DB 4; Length 270;
Best Local Similarity 66.7%; Pred. No. 89;
Matches              6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 LMAQEALAF 9
      |:| |||||
Db      90 MLAEALIF 98

RESULT 6
US-09-252-991A-26258
; Sequence 26258, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26258
; LENGTH: 503
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26258

Query Match          68.9%; Score 31; DB 4; Length 503;
Best Local Similarity 60.0%; Pred. No. 1.7e+02;
Matches              6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 LMAQEALAF 10
      |:| |||||
Db      244 LLAHALAYL 253

RESULT 7
US-09-198-452A-128
; Sequence 128, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragme
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, pr
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
```

; CURRENT APPLICATION NUMBER: US/09/198.452A  
; CURRENT FILING DATE: 1998-11-24  
; NUMBER OF SEQ ID NOS: 6849  
; SEQ ID NO 128  
; LENGTH: 636  
; TYPE: PRT  
; ORGANISM: Chlamydia pneumoniae  
US-09-198-452A-128

Query Match 68.9%; Score 31; DB 4; Length 636;  
Best Local Similarity 75.0%; Pred. No. 2.1e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MAQEALAF 9  
Db 58 LAQALAF 65

RESULT 8  
US-08-905-223-408  
; Sequence 408, Application US/08905223  
; Patent No. 6222029  
; GENERAL INFORMATION:  
; APPLICANT: Edwards, Jean-Baptiste D.  
; APPLICANT: Duclert, Aymeric  
; APPLICANT: Lacroix, Bruno  
; TITLE OF INVENTION: 5' ESTs FOR SECRETED PROTEINS  
; NUMBER OF SEQUENCES: 503  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson & Bear  
; STREET: 501 West Broadway  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92101-3505  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Win95  
; SOFTWARE: Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/905.223  
; FILING DATE:  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelsen, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 408:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 92 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PROTEIN  
; ORIGINAL SOURCE:  
; ORGANISM: Homo Sapiens  
; TISSUE TYPE: Brain  
; FEATURE:  
; NAME/KEY: sig\_peptide  
; LOCATION: -40...1  
; IDENTIFICATION METHOD: Von Heijne matrix  
; OTHER INFORMATION: score 7.9  
; OTHER INFORMATION: seq SLLLFQGFQPASS/QE

US-08-905-223-408  
Query Match 66.7%; Score 30; DB 3; Length 92;  
Best Local Similarity 77.8%; Pred. No. 47;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LMAQEALAF 9

Db 13 LMAQEHLEF 21

RESULT 9  
US-09-250-677-2  
; Sequence 2, Application US/09250677  
; Patent No. 6251631  
; GENERAL INFORMATION:  
; APPLICANT: Burnham, Martin Karl Russel  
; APPLICANT: Biswas, Sanjoy  
; APPLICANT: Chalker, Alison Francis  
; APPLICANT: Ingraham, Karen Anne  
; APPLICANT: Traini, Christopher Michael  
; APPLICANT: Warren, Patrick Vernon  
; TITLE OF INVENTION: nade  
; FILE REFERENCE: GM10197  
; CURRENT APPLICATION NUMBER: US/09/250.677  
; CURRENT FILING DATE: 1999-02-16  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 274  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-250-677-2

Query Match 66.7%; Score 30; DB 3; Length 274;  
Best Local Similarity 75.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 AQEALAF 10  
Db 94 AQKALAF 101

RESULT 10  
US-09-252-991A-28498  
; Sequence 28498, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252.991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 28498  
; LENGTH: 338  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-28498

Query Match 66.7%; Score 30; DB 4; Length 338;  
Best Local Similarity 60.0%; Pred. No. 1.8e+02;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LMAQEALAF 10  
Db 168 LFAQDALVIL 177

RESULT 11  
US-09-176-657-8  
; Sequence 8, Application US/09176657  
; Patent No. 6020164  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga

QY	DB	Score	DB 3;	Length	Indels	Mismatches	Gaps
1 LMAQEALA 8	249 LMAQSALA 256	66.7%	Score 30;	DB 3; Length 415;	0	0	0
<p>Query Match 66.7%; Score 30; DB 3; Length 415;</p> <p>Best Local Similarity 87.5%; Pred. No. 2.2e+02;</p> <p>Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;</p>							
<p>RESULT 12</p> <p>US-09-421-299-8</p> <p>Sequence 8, Application US/09421299</p> <p>Patent No. 6524579</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Bandman, Olga</p> <p>APPLICANT: Tang, Y. Tom</p> <p>APPLICANT: Corley, Neil C.</p> <p>APPLICANT: Guegler, Karl J.</p> <p>APPLICANT: Lu, Aina</p> <p>APPLICANT: Baughn, Mariah R.</p> <p>TITLE OF INVENTION: HUMAN RNA BINDING PROTEINS</p> <p>FILE REFERENCE: PF-0611 US</p> <p>CURRENT FILING DATE: 1999-10-20</p> <p>EARLIER FILING DATE: 1998-10-21</p> <p>NUMBER OF SEQ ID NOS: 9</p> <p>SOFTWARE: PERL Program</p> <p>SEQ ID NO 8</p> <p>LENGTH: 415</p> <p>TYPE: PRT</p> <p>ORGANISM: Caenorhabditis elegans</p> <p>FEATURE: -</p> <p>OTHER INFORMATION: g608464</p>							
1 LMAQEALA 8	249 LMAQSALA 256	66.7%	Score 30;	DB 3; Length 415;	0	0	0
<p>Query Match 66.7%; Score 30; DB 3; Length 415;</p> <p>Best Local Similarity 87.5%; Pred. No. 2.2e+02;</p> <p>Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;</p>							
<p>RESULT 13</p> <p>US-09-082-310-1</p> <p>Sequence 1, Application US/09082310</p> <p>Patent No. 6096526</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Hillman, Jennifer L.</p> <p>APPLICANT: Lal, Preeti</p> <p>APPLICANT: Corley, Neil C.</p> <p>APPLICANT: Guegler, Karl J.</p>							
1 LMAQEALA 8	249 LMAQSALA 256	66.7%	Score 30;	DB 4; Length 415;	0	0	0
<p>Query Match 66.7%; Score 30; DB 4; Length 415;</p> <p>Best Local Similarity 87.5%; Pred. No. 2.2e+02;</p> <p>Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;</p>							
<p>RESULT 14</p> <p>US-09-575-205-1</p> <p>Sequence 1, Application US/09575205</p> <p>Patent No. 6436683</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Hillman, Jennifer L.</p> <p>APPLICANT: Lal, Preeti</p> <p>APPLICANT: Corley, Neil C.</p> <p>APPLICANT: Guegler, Karl J.</p> <p>APPLICANT: Yue, Henry</p> <p>TITLE OF INVENTION: HUMAN NUCLEIC ACID METHYLASES</p> <p>NUMBER OF SEQUENCES: 4</p> <p>CORRESPONDENCE ADDRESS:</p> <p>ADDRESSEE: INCYTE PHARMACEUTICALS, INC.</p> <p>STREET: 3174 PORTER DRIVE</p> <p>CITY: PALO ALTO</p> <p>STATE: CALIFORNIA</p> <p>COUNTRY: USA</p> <p>ZIP: 94304</p> <p>COMPUTER READABLE FORM:</p> <p>MEDIUM TYPE: Floppy disk</p> <p>COMPUTER: IBM PC compatible</p> <p>OPERATING SYSTEM: PC-DOS/MS-DOS</p> <p>SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2</p> <p>CURRENT APPLICATION DATA:</p> <p>APPLICATION NUMBER: US/09/575,205</p> <p>FILING DATE:</p>							
1 LMAQEALAF 9	13 LMAQEHLEF 21	66.7%	Score 30;	DB 3; Length 77.8%;	0	0	0
<p>Query Match 66.7%; Score 30; DB 3; Length 77.8%;</p> <p>Best Local Similarity 77.8%; Pred. No. 2.4e+02;</p> <p>Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;</p>							
<p>US-09-082-310-1</p>							



```
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/082,310
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0520 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 463 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRSTNOT07
; CLONE: 2124957
; US-09-575-205-1
```

```
Query Match 66.7%; Score 30; DB 4; Length 463;
Best Local Similarity 77.8%; Pred. NO. 2.4e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 LMAQEALAF 9
    ||||| |
Db 13 LMAQEHLEF 21
```

```
RESULT 15
US-09-252-991A-30146
; Sequence 30146, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30146
; LENGTH: 763
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30146
```

```
Query Match 66.7%; Score 30; DB 4; Length 763;
Best Local Similarity 87.5%; Pred. NO. 4e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 3 AQEALAF 10
    |||||
Db 153 AQALAF 160
```

```
Search completed: October 7, 2003, 13:34:05
Job time : 5.44304 secs
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:31:20 ; Search time 10.8861 Seconds  
(without alignments)  
145.335 Million cell updates/sec

Title: US-09-807-512-12  
Perfect score: 45  
Sequence: 1 LMAQEAFL 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues  
Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	75.6	222	12	US-10-032-585-7802
2	33	73.3	891	10	US-09-949-192-45
3	32	71.1	204	15	US-10-156-761-13725
4	32	71.1	552	12	US-10-314-881-9
5	30	66.7	257	15	US-10-177-293-274
6	30	66.7	274	9	US-09-815-242-13286
7	30	66.7	274	10	US-09-855-400-2
8	30	66.7	415	12	US-10-359-385-8
9	30	66.7	463	14	US-10-155-613-1
10	30	66.7	1011	15	US-10-156-761-8999
11	29	64.4	29	12	US-10-289-135A-25
12	29	64.4	85	9	US-09-864-761-37705
13	29	64.4	153	9	US-09-939-980-507
14	29	64.4	249	9	US-09-205-658-309
15	29	64.4	249	12	US-09-963-693-309

16	29	64.4	305	12	US-10-219-810-38	Sequence 38, Appl
17	29	64.4	352	12	US-10-279-029-94	Sequence 94, Appl
18	29	64.4	352	12	US-10-219-810-21	Sequence 21, Appl
19	29	64.4	354	12	US-10-279-029-104	Sequence 104, App
20	29	64.4	354	12	US-10-219-810-31	Sequence 31, Appl
21	29	64.4	364	9	US-09-815-242-13652	Sequence 13652, A
22	29	64.4	394	12	US-10-299-003-17	Sequence 17, Appl
23	29	64.4	403	10	US-09-870-379-2	Sequence 2, Appl
24	29	64.4	403	12	US-10-299-003-2	Sequence 2, Appl
25	29	64.4	403	12	US-10-299-003-10	Sequence 10, Appl
26	29	64.4	403	12	US-10-299-003-49	Sequence 49, Appl
27	29	64.4	403	12	US-10-299-003-57	Sequence 57, Appl
28	29	64.4	403	12	US-10-299-003-57	Sequence 57, Appl
29	29	64.4	403	15	US-10-059-585-50	Sequence 50, Appl
30	29	64.4	424	15	US-10-156-761-14500	Sequence 14500, A
31	29	64.4	430	12	US-10-299-003-18	Sequence 18, Appl
32	29	64.4	559	12	US-10-205-219-36	Sequence 36, Appl
33	29	64.4	559	12	US-10-205-219-76	Sequence 76, Appl
34	29	64.4	559	12	US-10-299-003-15	Sequence 15, Appl
35	29	64.4	559	14	US-10-001-851-21	Sequence 21, Appl
36	29	64.4	597	12	US-10-299-003-8	Sequence 8, Appl
37	29	64.4	619	10	US-09-815-923-2	Sequence 2, Appl
38	29	64.4	734	14	US-10-008-355-5	Sequence 5, Appl
39	29	64.4	797	10	US-09-989-350-18	Sequence 18, Appl
40	28	62.2	18	15	US-10-084-813-267	Sequence 267, App
41	28	62.2	18	15	US-10-084-813-268	Sequence 268, App
42	28	62.2	86	15	US-10-156-761-14705	Sequence 14705, A
43	28	62.2	148	11	US-09-948-820-62	Sequence 62, Appl
44	28	62.2	267	15	US-10-156-761-14887	Sequence 14887, A
45	28	62.2	268	15	US-10-156-761-15002	Sequence 15002, A

ALIGNMENTS

RESULT 1

US-10-032-585-7802  
; Sequence 7802, Application US/10032585  
; Publication No. US20030180953A1  
; GENERAL INFORMATION:  
; APPLICANT: Terry, Roemer D.  
; APPLICANT: Bo, Jiang  
; APPLICANT: Charles, Boone  
; APPLICANT: Howard, Busone  
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery  
; FILE REFERENCE: 10182-005-999  
; CURRENT APPLICATION NUMBER: US/10/032,585  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 8000  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7802  
; LENGTH: 222  
; TYPE: PRT  
; ORGANISM: Candida albicans  
US-10-032-585-7802

Query Match 75.6%; Score 34; DB 12; Length 222;  
Best Local Similarity 80.0%; Pred No. 25;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMAQEAFL 10  
| | | | | | | | | |  
DB 122 LMAQEAFLV 131

RESULT 2

US-09-949-192-45  
; Sequence 45, Application US/09949192  
; Patent No. US20020142292A1  
; GENERAL INFORMATION:  
; APPLICANT: Parham, Christi L.  
; APPLICANT: Gorman, Daniel L.  
; APPLICANT: Kurata, Hirokazu

```
; APPLICANT: Arai, Naoko
; APPLICANT: Sana, Theodore R.
; APPLICANT: Mattson, Jeanine D.
; APPLICANT: Murphy, Erin E.
; APPLICANT: Savkoor, Chetan
; APPLICANT: Grein, Jeffery
; APPLICANT: Smith, Kathleen M.
; APPLICANT: McClanahan, Terrill K.
; TITLE OF INVENTION: MAMMALIAN GENES; RELATED REAGENTS AND METHODS
; FILE REFERENCE: DX01169X
; CURRENT APPLICATION NUMBER: US/09/949,192
; CURRENT FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/231,267
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 45
; LENGTH: 891
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-949-192-45

Query Match      73.3%; Score 33; DB 10; Length 891;
Best Local Similarity 87.5%; Pred. No. 1.9e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 AQEALAF 10
DB 130 SQEALAF 137

RESULT 3
US-10-156-761-13725
; Sequence 13725, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13725
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-13725

Query Match      71.1%; Score 32; DB 15; Length 204;
Best Local Similarity 70.0%; Pred. No. 59;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LMAQEALAF 10
DB 111 LVAQERAF 120

RESULT 4
US-10-314-881-9
; Sequence 9, Application US/10314881
; Publication No. US20030162279A1
; GENERAL INFORMATION:
; APPLICANT: Glucksman, Maria Alexandra
```

```
; APPLICANT: Williamson, Mark
; APPLICANT: Tsia, Fong-Ying
; APPLICANT: Rudolph-Owen, Laura A.
; TITLE OF INVENTION: 22438, 23553, 25278, and 26212 No. US20030162279A1e1
; TITLE OF INVENTION: human Sulfatases (A CIP Application)
; FILE REFERENCE: 35800/208398(5800-79
; CURRENT APPLICATION NUMBER: US/10/314,881
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US/09/773,426
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: US 09/495,823
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Pfam consensus sequence for human sulfatase
US-10-314-881-9

Query Match      71.1%; Score 32; DB 12; Length 552;
Best Local Similarity 70.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LMAQEALAF 10
DB 262 LLADEALPFL 271

RESULT 5
US-10-177-293-274
; Sequence 274, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
```

; SEQ ID NO 274  
; LENGTH: 257  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-177-293-274

Query Match 66.7%; Score 30; DB 15; Length 257;  
Best Local Similarity 77.8%; Pred. No. 2e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LMAQEALAF 9  
|||||  
Db 13 LMAQEHLEF 21

## RESULT 6

US-09-815-242-13286  
; Sequence 13286, Application US/09815242  
; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE OF INVENTION: Prokaryotes

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 13286

; LENGTH: 274

; TYPE: PRT

; ORGANISM: Streptococcus pneumoniae

US-09-815-242-13286

Query Match 66.7%; Score 30; DB 9; Length 274;  
Best Local Similarity 75.0%; Pred. No. 2.1e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 AQEALAF 10  
|||||  
Db 94 AQKALAFI 101

## RESULT 7

US-09-855-400-2

; Sequence 2, Application US/09855400

; Patent No. US20020115075A1

; GENERAL INFORMATION:

; APPLICANT: Burnham, Martin Karl Russel

; APPLICANT: Biswas, Sanjoy

; APPLICANT: Chalke, Alison Francis

; APPLICANT: Ingraham, Karen Anne

; APPLICANT: Traini, Christopher Michael  
; APPLICANT: Warren, Patrick Vernon  
; TITLE OF INVENTION: nade  
; FILE REFERENCE: GM10197  
; CURRENT APPLICATION NUMBER: US/09/855,400  
; CURRENT FILING DATE: 2001-05-15  
; PRIOR APPLICATION NUMBER: US/09/250,677  
; PRIOR FILING DATE: 1999-02-16  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 274  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-855-400-2

Query Match 66.7%; Score 30; DB 10; Length 274;  
Best Local Similarity 75.0%; Pred. No. 2.1e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 AQEALAF 10  
|||||  
Db 94 AQKALAFI 101

## RESULT 8

US-10-359-385-8

; Sequence 8, Application US/10359385

; Publication No. US20030143622A1

; GENERAL INFORMATION:

; APPLICANT: Bandman, Olga

; APPLICANT: Tang, Y. Tom

; APPLICANT: Corley, Neil C.

; APPLICANT: Guegler, Karl J.

; APPLICANT: Lu, Aina

; APPLICANT: Baughn, Mariah R.

; TITLE OF INVENTION: HUMAN RNA BINDING PROTEINS

; FILE REFERENCE: PF-0611 US

; CURRENT APPLICATION NUMBER: US/10/359,385

; CURRENT FILING DATE: 2003-02-05

; PRIOR APPLICATION NUMBER: US/09/176,657

; PRIOR FILING DATE: 1998-10-21

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PERL Program

; SEQ ID NO 8

; LENGTH: 415

; TYPE: PRT

; ORGANISM: Caenorhabditis elegans

; FEATURE: -

; OTHER INFORMATION: g608464

US-10-359-385-8

Query Match 66.7%; Score 30; DB 12; Length 415;  
Best Local Similarity 87.5%; Pred. No. 3.4e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMAQEALA 8  
|||||  
Db 249 LMAQSALA 256

## RESULT 9

US-10-155-613-1

; Sequence 1, Application US/10155613

; Publication No. US20020192787A1

; GENERAL INFORMATION:

; APPLICANT: Hillman, Jennifer L.

; APPLICANT: Lal, Preeti

; APPLICANT: Corley, Neil C.

; APPLICANT: Guegler, Karl J.

; APPLICANT: Yue, Henry

; TITLE OF INVENTION: HUMAN NUCLEIC ACID METHYLASES

; NUMBER OF SEQUENCES: 4

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/155,613
FILING DATE: 22-May-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/575,205
FILING DATE: <Unknown>
APPLICATION NUMBER: 09/082,310
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0520 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 463 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRSTNOT07
CLONE: 2124957
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-155-613-1
Query Match 66.7%; Score 30; DB 14; Length 463;
Best Local Similarity 77.8%; Pred. No. 3.8e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LMAQELAF 9
DB 13 LMAQEHLEF 21

RESULT 10
US-10-156-761-8999
; Sequence 8999, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 8999
; LENGTH: 1011
; TYPE: PRT

CORRESPONDENCE ADDRESS:
ADDRESSEE: Streptomyces avermitilis
US-10-156-761-8999
Query Match 66.7%; Score 30; DB 15; Length 1011;
Best Local Similarity 75.0%; Pred. No. 9e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 AQEALAF 10
DB 404 AEEAVAF 411

RESULT 11
US-10-289-135A-25
; Sequence 25, Application US/10289135A
; Publication No. US20030180937A1
; GENERAL INFORMATION:
; APPLICANT: GEORGIOU, GEORGE
; APPLICANT: DELISA, MATTHEW
; TITLE OF INVENTION: ENGINEERING OF LEADER PEPTIDES FOR THE SECRETION OF
; TITLE OF INVENTION: RECOMBINANT PROTEINS IN BACTERIA
; FILE REFERENCE: CLER:019US
; CURRENT APPLICATION NUMBER: US/10/289,135A
; CURRENT FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: 60/337,452
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-10-289-135A-25
Query Match 64.4%; Score 29; DB 12; Length 29;
Best Local Similarity 70.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 LMAQELAF 10
DB 12 LTASSALAF 21

RESULT 12
US-09-864-761-37705
; Sequence 37705, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
```

;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00664  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00669  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00665  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00668  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00663  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00662  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00661  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00670  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: US 60/234,687  
;; PRIOR FILING DATE: 2000-09-21  
;; PRIOR APPLICATION NUMBER: US 09/608,408  
;; PRIOR FILING DATE: 2000-06-30  
;; PRIOR APPLICATION NUMBER: US 09/774,203  
;; PRIOR FILING DATE: 2001-01-29  
;; NUMBER OF SEQ ID NOS: 49117  
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
;; SEQ ID NO 37705  
;; LENGTH: 85  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; OTHER INFORMATION: MAP TO AF067844.1  
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.5  
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.2  
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2  
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.4  
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.96  
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.9  
;; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.9  
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.5  
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2  
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2  
;; OTHER INFORMATION: SWISSPROT HIT: O08586, EVALUE 2.00e-44  
;; OTHER INFORMATION: EST\_HUMAN HIT: A1222037.1, EVALUE 2.00e-43  
US-09-864-761-37705

Query Match 64.4%; Score 29; DB 9; Length 85;  
Best Local Similarity 77.8%; Pred. No. 93;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LMAQEALAF 9  
| | | | |  
DB 67 LKAQEALDF 75

RESULT 13  
US-09-939-980-507  
; Sequence 507, Application us/09939980  
; Patent No. US20020082234A1  
; GENERAL INFORMATION:

;; APPLICANT: Black, Michael  
;; Burnham, Martin  
;; Hodgson, John  
;; Knowles, David  
;; Lonetto, Michael  
;; Nicholas, Michael  
;; Pratt, Julie  
;; Reichard, Richard  
;; Rosenberg, Martin  
;; Ward, Judith

;; TITLE OF INVENTION: Polypeptides and Their Uses  
;; NUMBER OF SEQUENCES: 534  
;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: SmithKline Beecham Corporation  
;; STREET: 709 Swedeland Road  
;; CITY: King of Prussia  
;; STATE: PA  
;; COUNTRY: USA  
;; ZIP: 19406-0939  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette  
;; COMPUTER: IBM Compatible  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: FastSeq for Windows Version 2.0  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/939,980  
;; FILING DATE: 27-Aug-2001  
;; CLASSIFICATION: <Unknown>  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/936,165  
;; FILING DATE: <Unknown>  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Gimmi, Edward R  
;; REGISTRATION NUMBER: 38,891  
;; REFERENCE/DOCKET NUMBER: P50549  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 610-270-4478  
;; TELEFAX: 610-270-5090  
;; TELEX: <Unknown>  
;; INFORMATION FOR SEQ ID NO: 507:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 153 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: Protein  
;; SEQUENCE DESCRIPTION: SEQ ID NO: 507:  
US-09-939-980-507

Query Match 64.4%; Score 29; DB 9; Length 153;  
Best Local Similarity 66.7%; Pred. No. 1.8e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMAQEALAF 9  
| | | | |  
DB 69 LTAEEALSF 77

RESULT 14  
US-09-203-658-309  
; Sequence 309, Application us/09205658  
; Patent No. US20010029617A1  
; GENERAL INFORMATION:

;; APPLICANT: Ruvkun, Gary  
;; APPLICANT: Ogg, Scott  
;; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR  
;; TITLE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS  
;; FILE REFERENCE: 00786/351004  
;; CURRENT APPLICATION NUMBER: US/09/205,658  
;; CURRENT FILING DATE: 1998-12-03  
;; EARLIER APPLICATION NUMBER: 08/857,076  
;; EARLIER FILING DATE: 1997-05-15  
;; EARLIER APPLICATION NUMBER: 08/888,534  
;; EARLIER FILING DATE: 1997-07-07  
;; EARLIER APPLICATION NUMBER: US98/10080  
;; EARLIER FILING DATE: 1998-05-15  
;; NUMBER OF SEQ ID NOS: 328  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 309  
;; LENGTH: 249  
;; TYPE: PRT  
;; ORGANISM: Caenorhabditis elegans

Query Match 64.4%; Score 29; DB 9; Length 249;  
Best Local Similarity 77.8%; Pred. No. 3.1e+02;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LMAQALAF 9  
| | | | |  
Db 143 LKAEALDF 151

RESULT 15

US-09-963-693-309  
; Sequence 309, Application US/09963693  
; Publication No. US20030181364A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruvkun, Gary  
; APPLICANT: O'Connell, Scott  
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR  
; TITLE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS  
; FILE REFERENCE: 00786/351004  
; CURRENT APPLICATION NUMBER: US/09/963,693  
; CURRENT FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US/09/205,658  
; PRIOR FILING DATE: 1998-12-03  
; PRIOR APPLICATION NUMBER: 08/857,076  
; PRIOR FILING DATE: 1997-05-15  
; PRIOR APPLICATION NUMBER: 08/888,534  
; PRIOR FILING DATE: 1997-07-07  
; PRIOR APPLICATION NUMBER: US98/10080  
; PRIOR FILING DATE: 1998-05-15  
; NUMBER OF SEQ ID NOS: 328  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 309  
; LENGTH: 249  
; TYPE: PRT  
; ORGANISM: Caenorhabditis elegans  
US-09-963-693-309

Query Match 64.4%; Score 29; DB 12; Length 249;  
Best Local Similarity 77.8%; Pred. No. 3.1e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LMAQALAF 9  
| | | | |  
Db 143 LKAEALDF 151

Search completed: October 7, 2003, 13:58:55  
Job time : 10.8861 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:25:43 ; Search time 78.7342 Seconds  
(without alignments)  
115.568 Million cell updates/sec

Title: US-09-807-512-12  
Perfect score: 45  
Sequence: 1 LMAQEAFL 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues  
Total number of hits satisfying chosen parameters: 5728757

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main:\*

1:	/cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
2:	/cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3:	/cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4:	/cgn2_6/ptodata/1/paa/US080_COMB.pep.*
5:	/cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6:	/cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7:	/cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8:	/cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9:	/cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10:	/cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11:	/cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12:	/cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13:	/cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14:	/cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15:	/cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16:	/cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17:	/cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18:	/cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19:	/cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20:	/cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21:	/cgn2_6/ptodata/1/paa/US097A_COMB.pep.*
22:	/cgn2_6/ptodata/1/paa/US097B_COMB.pep.*
23:	/cgn2_6/ptodata/1/paa/US098_COMB.pep.*
24:	/cgn2_6/ptodata/1/paa/US099A_COMB.pep.*
25:	/cgn2_6/ptodata/1/paa/US099B_COMB.pep.*
26:	/cgn2_6/ptodata/1/paa/US100_COMB.pep.*
27:	/cgn2_6/ptodata/1/paa/US101_COMB.pep.*
28:	/cgn2_6/ptodata/1/paa/US102_COMB.pep.*
29:	/cgn2_6/ptodata/1/paa/US103_COMB.pep.*
30:	/cgn2_6/ptodata/1/paa/US104_COMB.pep.*
31:	/cgn2_6/ptodata/1/paa/US106_COMB.pep.*
32:	/cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	100.0	10	23	US-09-807-512-12
					Sequence 12, Appl

2	45	100.0	11	1	PCT-US03-16736-81	Sequence 81, Appl
3	45	100.0	11	1	PCT-US03-16736-82	Sequence 82, Appl
4	45	100.0	11	23	US-09-807-512-11	Sequence 11, Appl
5	45	100.0	11	30	US-10-447-161-81	Sequence 81, Appl
6	45	100.0	11	30	US-10-447-161-82	Sequence 82, Appl
7	45	100.0	25	19	US-09-529-206-56	Sequence 56, Appl
8	45	100.0	25	19	US-09-529-206B-56	Sequence 56, Appl
9	45	100.0	25	19	US-09-529-206D-56	Sequence 56, Appl
10	45	100.0	25	19	US-09-529-206-5	Sequence 5, Appl
11	45	100.0	58	19	US-09-529-206A-5	Sequence 5, Appl
12	45	100.0	58	19	US-09-529-206B-5	Sequence 5, Appl
13	45	100.0	58	19	US-09-529-206D-5	Sequence 5, Appl
14	45	100.0	58	23	US-09-807-512-10	Sequence 10, Appl
15	45	100.0	109	23	US-09-807-512-2	Sequence 2, Appl
16	45	100.0	218	24	US-09-902-540-14528	Sequence 14528, A
17	36	80.0	513	15	US-09-134-000-5326	Sequence 5326, Ap
18	36	80.0	513	15	US-09-134-000C-5326	Sequence 5326, Ap
19	36	80.0	513	30	US-10-434-665-5326	Sequence 5326, Ap
20	36	80.0	222	1	PCT-US02-03987-15746	Sequence 15746, A
21	34	75.6	222	26	US-10-032-585-7802	Sequence 7802, Ap
22	34	75.6	222	26	US-10-072-851-15746	Sequence 15746, A
23	34	75.6	222	32	US-60-314-050-7802	Sequence 7802, Ap
24	34	75.6	232	16	US-09-248-796-15087	Sequence 15087, A
25	34	75.6	232	31	US-10-603-113-15087	Sequence 15087, A
26	34	75.6	232	32	US-60-096-409-15087	Sequence 15087, A
27	34	75.6	484	28	US-10-282-122A-47591	Sequence 47591, A
28	34	75.6	577	26	US-10-092-900A-338	Sequence 338, App
29	34	75.6	115	20	US-09-620-393B-2247	Sequence 2247, Ap
30	33	73.3	320	19	US-09-513-996A-24785	Sequence 24785, A
31	33	73.3	333	19	US-09-573-655A-886	Sequence 886, App
32	33	73.3	333	19	US-09-573-655A-1148	Sequence 1148, Ap
33	33	73.3	333	19	US-09-573-655B-886	Sequence 886, App
34	33	73.3	333	19	US-09-573-655B-1148	Sequence 1148, Ap
35	33	73.3	333	21	US-09-708-427-24479	Sequence 24479, A
36	33	73.3	761	26	US-10-070-226B-15	Sequence 15, Appl
37	33	73.3	891	1	PCT-US02-30474-1070	Sequence 1070, Ap
38	33	73.3	891	1	PCT-US02-35563-2	Sequence 2, Appl
39	33	73.3	891	1	PCT-US03-14742-12	Sequence 12, Appl
40	33	73.3	891	24	US-09-949-132-45	Sequence 45, Appl
41	33	73.3	891	32	US-60-324-631-1075	Sequence 1075, Ap
42	33	73.3	128	1	PCT-US02-36123-4400	Sequence 4400, Ap
43	32	71.1	204	27	US-10-156-761-13725	Sequence 13725, A
44	32	71.1				
45						

ALIGNMENTS

RESULT 1  
US-09-807-512-12  
; Sequence 12, Application US/09807512  
; GENERAL INFORMATION:  
; APPLICANT: Schrier, Peter I.  
; APPLICANT: Aarnoudse, Corlien  
; APPLICANT: Heider, Karl-Heinz  
; APPLICANT: Klade, Christoph  
; TITLE OF INVENTION: Camel, An Alternative Translation Product of the Tumor Antigen-Lage 1  
; FILE REFERENCE: 0652.220000  
; CURRENT APPLICATION NUMBER: US/09/807,512  
; CURRENT FILING DATE: 2001-04-16  
; PRIOR APPLICATION NUMBER: PCT/EP99/07832  
; PRIOR FILING DATE: 1999-10-15  
; PRIOR APPLICATION NUMBER: EP 98119583.7  
; PRIOR FILING DATE: 1998-10-16  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 12  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-807-512-12

Query Match 100.0%; Score 45; DB 23; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.039;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMAQEALAF 10  
 |||||  
 Db 1 LMAQEALAF 10

## RESULT 2

PCT-US03-16736-81

; Sequence 81, Application PC/TUS0316736  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Rong-fu  
 ; TITLE OF INVENTION: Mutant Fibronectin and Tumor Metastasis  
 ; FILE REFERENCE: HO-P02484W00  
 ; CURRENT APPLICATION NUMBER: PCT/US03/16736  
 ; CURRENT FILING DATE: 2003-05-28  
 ; PRIOR APPLICATION NUMBER: 60/383,530  
 ; PRIOR FILING DATE: 2002-05-28  
 ; NUMBER OF SEQ ID NOS: 148  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 81  
 ; LENGTH: 11  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Synthetic Peptide  
 PCT-US03-16736-81

Query Match 100.0%; Score 45; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.043;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMAQEALAF 10  
 |||||  
 Db 2 LMAQEALAF 11

## RESULT 3

PCT-US03-16736-82

; Sequence 82, Application PC/TUS0316736  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Rong-fu  
 ; TITLE OF INVENTION: Mutant Fibronectin and Tumor Metastasis  
 ; FILE REFERENCE: HO-P02484W00  
 ; CURRENT APPLICATION NUMBER: PCT/US03/16736  
 ; CURRENT FILING DATE: 2003-05-28  
 ; PRIOR APPLICATION NUMBER: 60/383,530  
 ; PRIOR FILING DATE: 2002-05-28  
 ; NUMBER OF SEQ ID NOS: 148  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 82  
 ; LENGTH: 11  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Synthetic Peptide  
 PCT-US03-16736-82

Query Match 100.0%; Score 45; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.043;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMAQEALAF 10  
 |||||  
 Db 2 LMAQEALAF 11

## RESULT 4

US-09-807-512-11  
 ; Sequence 11, Application US/09807512

; GENERAL INFORMATION:  
 ; APPLICANT: Schrier, Peter I.  
 ; APPLICANT: Aarnoudse, Corlien  
 ; APPLICANT: Heider, Karl-Heinz  
 ; APPLICANT: Klade, Christoph  
 ; TITLE OF INVENTION: Camel, An Alternative Translation Product of the Tumor  
 ; TITLE OF INVENTION: Antigen-Lage 1  
 ; FILE REFERENCE: 0652.2200000  
 ; CURRENT APPLICATION NUMBER: US/09/807,512  
 ; CURRENT FILING DATE: 2001-04-16  
 ; PRIOR APPLICATION NUMBER: PCT/EP99/07832  
 ; PRIOR FILING DATE: 1999-10-15  
 ; PRIOR APPLICATION NUMBER: EP 98119583.7  
 ; PRIOR FILING DATE: 1998-10-16  
 ; NUMBER OF SEQ ID NOS: 28  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 11  
 ; LENGTH: 11  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; OTHER INFORMATION: Synthetic Peptide  
 US-09-807-512-11

Query Match 100.0%; Score 45; DB 23; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.043;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMAQEALAF 10  
 |||||  
 Db 2 LMAQEALAF 11

## RESULT 5

US-10-447-161-81

; Sequence 81, Application US/10447161  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Rong-fu  
 ; TITLE OF INVENTION: Mutant Fibronectin and Tumor Metastasis  
 ; FILE REFERENCE: HO-P02484US1  
 ; CURRENT APPLICATION NUMBER: US/10/447,161  
 ; CURRENT FILING DATE: 2003-05-28  
 ; PRIOR APPLICATION NUMBER: 60/383,530  
 ; PRIOR FILING DATE: 2002-05-28  
 ; NUMBER OF SEQ ID NOS: 148  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 81  
 ; LENGTH: 11  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Synthetic Peptide  
 US-10-447-161-81

Query Match 100.0%; Score 45; DB 30; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.043; 0; Indels 0; Gaps 0;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMAQEALAF 10  
 |||||  
 Db 2 LMAQEALAF 11

## RESULT 6

US-10-447-161-82

; Sequence 82, Application US/10447161  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Rong-fu  
 ; TITLE OF INVENTION: Mutant Fibronectin and Tumor Metastasis  
 ; FILE REFERENCE: HO-P02484US1  
 ; CURRENT APPLICATION NUMBER: US/10/447,161  
 ; CURRENT FILING DATE: 2003-05-28  
 ; PRIOR APPLICATION NUMBER: 60/383,530  
 ; PRIOR FILING DATE: 2002-05-28  
 ; NUMBER OF SEQ ID NOS: 148

; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 82  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Peptide  
US-10-447-161-82

Query Match 100.0%; Score 45; DB 30; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.043; 0; Indels 0; Gaps 0;  
Matches 10; Conservative 0; Mismatches 0;

QY 1 LMAQEALAF 10  
| | | | | | | | | |  
Db 2 LMAQEALAF 11

RESULT 7  
US-09-529-206-56  
; Sequence 56, Application US/09529206  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America as  
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene  
; FILE REFERENCE: 20264269PC  
; CURRENT APPLICATION NUMBER: US/09/529,206  
; PRIOR FILING DATE: 2000-06-13  
; PRIOR APPLICATION NUMBER: US60/061,428  
; PRIOR FILING DATE: 1997-10-08  
; NUMBER OF SEQ ID NOS: 106  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 56  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-206-56

Query Match 100.0%; Score 45; DB 19; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.11; 0; Indels 0; Gaps 0;  
Matches 10; Conservative 0; Mismatches 0;

QY 1 LMAQEALAF 10  
| | | | | | | | | |  
Db 2 LMAQEALAF 11

RESULT 8  
US-09-529-206A-56  
; Sequence 56, Application US/09529206A  
; GENERAL INFORMATION:  
; APPLICANT: Wang Rong, Fu  
; APPLICANT: Rosenberg, Steven  
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene  
; FILE REFERENCE: 20264269US1  
; CURRENT APPLICATION NUMBER: US/09/529,206A  
; PRIOR FILING DATE: 2000-06-13  
; PRIOR APPLICATION NUMBER: PCT/US98/19609  
; PRIOR FILING DATE: 1998-09-21  
; PRIOR APPLICATION NUMBER: US60/061,428  
; PRIOR FILING DATE: 1997-10-08  
; NUMBER OF SEQ ID NOS: 127  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 56  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-206A-56

Query Match 100.0%; Score 45; DB 19; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.11; 0; Indels 0; Gaps 0;  
Matches 10; Conservative 0; Mismatches 0;

QY 1 LMAQEALAF 10  
| | | | | | | | | |  
Db 2 LMAQEALAF 11

RESULT 9  
US-09-529-206B-56  
; Sequence 56, Application US/09529206B  
; GENERAL INFORMATION:  
; APPLICANT: Wang Rong, Fu  
; APPLICANT: Rosenberg, Steven  
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene  
; FILE REFERENCE: 20264269US1  
; CURRENT APPLICATION NUMBER: US/09/529,206B  
; PRIOR FILING DATE: 2000-06-13  
; PRIOR APPLICATION NUMBER: PCT/US98/19609  
; PRIOR FILING DATE: 1998-09-21  
; PRIOR APPLICATION NUMBER: US60/061,428  
; PRIOR FILING DATE: 1997-10-08  
; NUMBER OF SEQ ID NOS: 127  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 56  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-206B-56

Query Match 100.0%; Score 45; DB 19; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.11; 0; Indels 0; Gaps 0;  
Matches 10; Conservative 0; Mismatches 0;

QY 1 LMAQEALAF 10  
| | | | | | | | | |  
Db 2 LMAQEALAF 11

RESULT 10  
US-09-529-206D-56  
; Sequence 56, Application US/09529206D  
; GENERAL INFORMATION:  
; APPLICANT: Wang Rong, Fu  
; APPLICANT: Rosenberg, Steven  
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene  
; FILE REFERENCE: 218791  
; CURRENT APPLICATION NUMBER: US/09/529,206D  
; CURRENT FILING DATE: 2000-06-13  
; PRIOR APPLICATION NUMBER: PCT/US98/19609  
; PRIOR FILING DATE: 1998-09-21  
; PRIOR APPLICATION NUMBER: US60/061,428  
; PRIOR FILING DATE: 1997-10-08  
; NUMBER OF SEQ ID NOS: 127  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 56  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-206D-56

Query Match 100.0%; Score 45; DB 19; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.11; 0; Indels 0; Gaps 0;  
Matches 10; Conservative 0; Mismatches 0;

QY 1 LMAQEALAF 10  
| | | | | | | | | |  
Db 2 LMAQEALAF 11

RESULT 11  
US-09-529-206-5  
; Sequence 5, Application US/09529206

Tue Oct 7 15:52:24 2003

us-09-807-512-12.rapm

```

; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America as
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; TITLE OF INVENTION: Encoding Same
; FILE REFERENCE: 20264269PC
; CURRENT APPLICATION NUMBER: US/09/529,206
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206-5

Query Match 100.0%; Score 45; DB 19; Length 58;
Best Local Similarity 100.0%; Pred. No. 0.29; 0; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0;

QY 1 LMAQEALAF 10
Db 2 LMAQEALAF 11

RESULT 12
US-09-529-206A-5
; Sequence 5, Application US/09529206A
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; TITLE OF INVENTION: Encoding Same
; FILE REFERENCE: 20264269US1
; CURRENT APPLICATION NUMBER: US/09/529,206A
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: PCT/US98/19609
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206A-5

Query Match 100.0%; Score 45; DB 19; Length 58;
Best Local Similarity 100.0%; Pred. No. 0.29; 0; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0;

QY 1 LMAQEALAF 10
Db 2 LMAQEALAF 11

RESULT 13
US-09-529-206B-5
; Sequence 5, Application US/09529206B
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; TITLE OF INVENTION: Encoding Same
; FILE REFERENCE: 20264269US1
; CURRENT APPLICATION NUMBER: US/09/529,206B
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: PCT/US98/19609
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08

```

```

; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206B-5

Query Match 100.0%; Score 45; DB 19; Length 58;
Best Local Similarity 100.0%; Pred. No. 0.29; 0; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0;

QY 1 LMAQEALAF 10
Db 2 LMAQEALAF 11

RESULT 14
US-09-529-206D-5
; Sequence 5, Application US/09529206D
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; TITLE OF INVENTION: Encoding Same
; FILE REFERENCE: 218791
; CURRENT APPLICATION NUMBER: US/09/529,206D
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: PCT/US98/19609
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206D-5

Query Match 100.0%; Score 45; DB 19; Length 58;
Best Local Similarity 100.0%; Pred. No. 0.29; 0; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0;

QY 1 LMAQEALAF 10
Db 2 LMAQEALAF 11

RESULT 15
US-09-807-512-10
; Sequence 10, Application US/09807512
; GENERAL INFORMATION:
; APPLICANT: Schrier, Peter I.
; APPLICANT: Aarnoudse, Corlien
; APPLICANT: Heider, Karl-Heinz
; APPLICANT: Klade, Christoph
; TITLE OF INVENTION: Camel, An Alternative Translation Product of the Tumor
; TITLE OF INVENTION: Antigen-Lage 1
; FILE REFERENCE: 0652.2200000
; CURRENT APPLICATION NUMBER: US/09/807,512
; CURRENT FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: PCT/EP99/07832
; PRIOR FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: EP 98119583.7
; PRIOR FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-807-512-10

```

Query Match 100.0%; Score 45; DB 23; Length 58;  
Best Local Similarity 100.0%; Pred. No. 0.29;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMAQEALAF 10  
Db 2 LMAQEALAF 11

Search completed: October 7, 2003, 13:54:59  
Job time : 79.7342 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:26:33 ; Search time 3.16456 Seconds  
(without alignments)  
92.039 Million cell updates/sec

Title: US-09-807-512-12

Perfect score: 45

Sequence: 1 LMAQEAALFL 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 136899 seqs, 29126274 residues

Total number of hits satisfying chosen parameters: 136899

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_New:\*

1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep:\*

2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep:\*

3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep:\*

4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep:\*

5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep:\*

6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:\*

7: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	30	6	US-10-296-734-1422
2	45	100.0	58	6	US-10-296-734-833
3	45	100.0	3541	6	US-10-296-734-1454
4	31	68.9	225	6	US-10-425-114A-46642
5	31	68.9	982	7	US-60-495-114-1573
6	31	68.9	2347	5	US-09-897-516A-7998
7	31	68.9	2538	1	PCT-US03-20082-49
8	30	66.7	346	6	US-10-425-114A-49166
9	30	66.7	349	6	US-10-425-114A-48214
10	30	66.7	445	1	PCT-US03-28227-5428
11	30	66.7	445	1	PCT-US03-28227-5429
12	30	66.7	445	1	PCT-US03-28227-5430
13	30	66.7	445	1	PCT-US03-28227-5431
14	29	64.4	29	1	PCT-US02-35618B-25
15	29	64.4	135	6	US-10-425-114A-38464
16	29	64.4	179	6	US-10-425-114A-44651
17	29	64.4	241	5	US-09-897-516A-5870
18	29	64.4	249	5	US-09-963-693A-309
19	29	64.4	364	6	US-10-425-114A-43937
20	29	64.4	399	6	US-10-425-114A-43897
21	29	64.4	399	6	US-10-425-114A-43914
22	29	64.4	399	6	US-10-425-114A-43996
23	29	64.4	399	6	US-10-425-114A-43996
24	29	64.4	399	6	US-10-425-114A-45802
25	29	64.4	399	6	US-10-425-114A-45873
26	29	64.4	402	6	US-10-425-114A-63733

## ALIGNMENTS

## RESULT 1

US-10-296-734-1422

; Sequence 1422, Application US/10296734

; GENERAL INFORMATION:

; APPLICANT: Thompson, Scott A

; APPLICANT: Ramshaw, Ian A

; TITLE OF INVENTION: Synthetic molecules and uses therefor

; FILE REFERENCE: Savine

; CURRENT APPLICATION NUMBER: US/10/296,734

; CURRENT FILING DATE: 2003-08-04

; PRIOR APPLICATION NUMBER: AU PQ7761/00

; PRIOR FILING DATE: 2000-05-26

; NUMBER OF SEQ ID NOS: 1507

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 1422

; LENGTH: 30

; TYPE: PRT

; ORGANISM: Artificial

; FEATURE:

; OTHER INFORMATION: NYsolb segment 1

US-10-296-734-1422

Query Match 100.0%; Score 45; DB 6; Length 30;

Best Local Similarity 100.0%; Pred. No. 0.1;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LMAQEAALFL 10

Db 4 LMAQEAALFL 13

## RESULT 2

US-10-296-734-833

; Sequence 833, Application US/10296734

; GENERAL INFORMATION:

; APPLICANT: Thompson, Scott A

; APPLICANT: Ramshaw, Ian A

; TITLE OF INVENTION: Synthetic molecules and uses therefor

; FILE REFERENCE: Savine

; CURRENT APPLICATION NUMBER: US/10/296,734

; CURRENT FILING DATE: 2003-08-04

; PRIOR APPLICATION NUMBER: AU PQ7761/00

; PRIOR FILING DATE: 2000-05-26

; NUMBER OF SEQ ID NOS: 1507

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 833

; LENGTH: 58

; TYPE: PRT

; ORGANISM: Artificial

Sequence 63734, A  
Sequence 42053, A  
Sequence 57203, A  
Sequence 109, App  
Sequence 109, App  
Sequence 1062, App  
Sequence 1063, App  
Sequence 47493, A  
Sequence 563, App  
Sequence 7176, App  
Sequence 41954, A  
Sequence 44, Appl  
Sequence 1065, App  
Sequence 2, Appl  
Sequence 40320, A  
Sequence 42082, A  
Sequence 45530, A  
Sequence 39375, A  
Sequence 493, App

```
; FEATURE:
; OTHER INFORMATION: NYN501b consensus polypeptide
US-10-296-734-833

Query Match      100.0%; Score 45; DB 6; Length 58;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMAQEALAF 10
    |||||
Db 2 LMAQEALAF 11
    |||||

RESULT 3
US-10-296-734-1454
; Sequence 1454, Application US/10296734
; GENERAL INFORMATION:
; APPLICANT: Thompson, Scott A
; TITLE OF INVENTION: Synthetic molecules and uses therefor
; FILE REFERENCE: Savine
; CURRENT APPLICATION NUMBER: US/10/296,734
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: AU P07761/00
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1454
; LENGTH: 3541
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Melanoma cancer specific savine
US-10-296-734-1454

Query Match      100.0%; Score 45; DB 6; Length 3541;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMAQEALAF 10
    |||||
Db 1755 LMAQEALAF 1764
    |||||

RESULT 4
US-10-425-114A-46642
; Sequence 46642, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 46642
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 701150419_FLI.pep
US-10-425-114A-46642

Query Match      68.9%; Score 31; DB 6; Length 225;
Best Local Similarity 77.8%; Pred. No. 86;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 MAQEALAF 10
    |||||
```

```
Db 112 MAQEADAF 120
    |||||

RESULT 5
US-60-495-114-1573
; Sequence 1573, Application US/60495114
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; FILE REFERENCE: ENCODING HUMAN PROTEASE PROTEINS, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001480
; CURRENT APPLICATION NUMBER: US/60/495,114
; CURRENT FILING DATE: 2003-08-15
; NUMBER OF SEQ ID NOS: 91238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1573
; LENGTH: 982
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-495-114-1573

Query Match      68.9%; Score 31; DB 7; Length 982;
Best Local Similarity 87.5%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 AQEALAF 10
    |||||
Db 205 AQELLAF 212
    |||||

RESULT 6
US-09-897-516A-7998
; Sequence 7998, Application US/09897516A
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesling, Joseph E.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8415
; SEQ ID NO 7998
; LENGTH: 2347
; TYPE: PRT
; ORGANISM: Xenorhabdus sp.
US-09-897-516A-7998

Query Match      68.9%; Score 31; DB 5; Length 2347;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMAQEAL 7
    |||||
Db 2290 LMAQEAL 2296
    |||||

RESULT 7
PCT-US03-20082-49
; Sequence 49, Application PC/TUS0320082
; GENERAL INFORMATION:
; APPLICANT: Dow AgroSciences LLC
; TITLE OF INVENTION: Pesticidally Active Proteins and Polynucleotides Obtainable fr
; TITLE OF INVENTION: Paenibacillus Species
```



FILE REFERENCE: DAS-101XC2  
CURRENT APPLICATION NUMBER: PCT/US03/20082  
CURRENT FILING DATE: 2003-06-27  
PRIOR APPLICATION NUMBER: US 60/392,633  
PRIOR FILING DATE: 2002-06-28  
PRIOR APPLICATION NUMBER: US 60/441,647  
PRIOR FILING DATE: 2003-01-21  
NUMBER OF SEQ ID NOS: 49  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 49  
LENGTH: 2538  
TYPE: PRT  
ORGANISM: Xenorhabdus nematophilus  
PCT-US03-20082-49

Query Match 68.9%; Score 31; DB 1; Length 2538;  
Best Local Similarity 100.0%; Pred. No. 4.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMAQEAFL 7  
| | | | |  
Db 2305 LMAQEAFL 2311

## RESULT 8

US-10-425-114A-49166  
Sequence 49166, Application US/10425114A  
GENERAL INFORMATION:  
APPLICANT: Liu, Jingdong  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K.  
APPLICANT: Screen, Steven E.  
APPLICANT: Tabaska, Jack E.  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53313)B  
CURRENT APPLICATION NUMBER: US/10/425,114A  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 49166  
LENGTH: 346  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: UC-ZMFLM017085B01\_FLI.pep  
US-10-425-114A-49166

Query Match 66.7%; Score 30; DB 6; Length 346;  
Best Local Similarity 70.0%; Pred. No. 1.7e+02;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 LMAQEAFL 10  
| | | | |  
Db 177 LMQEQLAFL 186

## RESULT 9

US-10-425-114A-48214  
Sequence 48214, Application US/10425114A  
GENERAL INFORMATION:  
APPLICANT: Liu, Jingdong  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K.  
APPLICANT: Screen, Steven E.  
APPLICANT: Tabaska, Jack E.  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53313)B  
CURRENT APPLICATION NUMBER: US/10/425,114A  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128

SEQ ID NO 48214  
LENGTH: 349  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: LIB3595-027-cl0\_FLI.pep  
US-10-425-114A-48214

Query Match 66.7%; Score 30; DB 6; Length 349;  
Best Local Similarity 70.0%; Pred. No. 1.7e+02;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 LMAQEAFL 10  
| | | | |  
Db 177 LMQEQLAFL 186

## RESULT 10

PCT-US03-28227-5428  
Sequence 5428, Application PC/TUS0328227  
GENERAL INFORMATION:  
APPLICANT: INCYTE CORPORATION; SCHMIDT, Jeanette P.;  
APPLICANT: WRIGHT, Rachel J.; BRUNS, Christopher M.;  
APPLICANT: MARJANOVIC, Mirjana M.; SHEN, Fan;  
APPLICANT: HARTSHORNE, Toinette A.; SUCHOROLSKI, Martin;  
APPLICANT: ALTUS, Christina M.; PITTS, Steven J.;  
APPLICANT: ELDER, Linda V.; MOONEY, Elizabeth M.;  
APPLICANT: DELEGEANE, Angelo M.; PANESAR, Iqbal S.;  
APPLICANT: BANVILLE, Steven C.; REDDY, Thirupathi P.;  
APPLICANT: STEVENS, Kristian A.; BLANCHARD, John L.;  
APPLICANT: PANZER, Scott R.; WANG, Xinhao;  
APPLICANT: AU, Alan P.; GERSTIN, Edward H., Jr.;  
APPLICANT: PERALTA, Careyna H.; ANDERSON, Scott E.;  
APPLICANT: RIOUX, Pierre; SHEN, Edward J.;  
APPLICANT: WU, Mingham C.; STUVE, Laura L.;  
APPLICANT: LAGACE, Robert E.; SPIRO, Peter A.;  
APPLICANT: STEWART, Elizabeth A.; WINGROVE, James A.;  
APPLICANT: VITT, Ursula A.; KIRTON, Edward;  
APPLICANT: XU, Yuming; KWONG, Mary;  
APPLICANT: POLICKY, Jennifer L.; HURWITZ, Bonnie L.;  
APPLICANT: MA, Yan; JACKSON, Jennifer L.;  
APPLICANT: GIETZEN, Darryl; PATURY, Srikanth;  
APPLICANT: SHI, Xiaobing; SUAREZ, Charlyn J.  
TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: PN-0100 PCT  
CURRENT APPLICATION NUMBER: PCT/US03/28227  
CURRENT FILING DATE: 2003-09-12  
PRIOR APPLICATION NUMBER: US 60/410,260  
PRIOR FILING DATE: 2002-09-12  
PRIOR APPLICATION NUMBER: US 60/410,259  
PRIOR FILING DATE: 2002-09-12  
NUMBER OF SEQ ID NOS: 5444  
SOFTWARE: PERL Program  
SEQ ID NO 5428  
LENGTH: 445  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc.feature  
OTHER INFORMATION: Incyte ID No: 966243.PT6p  
PCT-US03-28227-5428

Query Match 66.7%; Score 30; DB 1; Length 445;  
Best Local Similarity 50.0%; Pred. No. 2e+02;  
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMAQEAFL 10  
| | | | |  
Db 266 IMVKEAISFL 275

## RESULT 11

PCT-US03-28227-5429

```

; Sequence 5429, Application PC/TUS0328227
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; SCHMIDT, Jeanette P.;
; APPLICANT: WRIGHT, Rachel J.; BRUNS, Christopher M.;
; APPLICANT: MARJANOVIC, Mirjana M.; SHEN, Fan;
; APPLICANT: HARTSHORNE, Toinette A.; SUCHOROLSKI, Martin;
; APPLICANT: ALTUS, Christina M.; PITTS, Steven J.;
; APPLICANT: ELDER, Linda V.; MOONEY, Elizabeth M.;
; APPLICANT: DELEGEANE, Angelo M.; PANESAR, Iqbal S.;
; APPLICANT: BANVILLE, Steven C.; REDDY, Thirupathi P.;
; APPLICANT: STEVENS, Kristian A.; BLANCHARD, John L.;
; APPLICANT: PANZER, Scott R.; WANG, Xinhao;
; APPLICANT: AU, Alan P.; GERSTIN, Edward H., Jr.;
; APPLICANT: PERALTA, Careyna H.; ANDERSON, Scott E.;
; APPLICANT: WU, Mingham C.; STUVE, Laura L.;
; APPLICANT: LAGACE, Robert E.; SPIRO, Peter A.;
; APPLICANT: STEWART, Elizabeth A.; WINGROVE, James A.;
; APPLICANT: VITT, Ursula A.; KIRTON, Edward;
; APPLICANT: POLICKY, Jennifer L.; HURWITZ, Bonnie L.;
; APPLICANT: MA, Yan; JACKSON, Jennifer L.;
; APPLICANT: GIETZEN, Darryl; PATURY, Srikanth;
; APPLICANT: SHI, Xiaobing; SUAREZ, Charlyn J.
; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PN-0100 PCT
; CURRENT APPLICATION NUMBER: US 60/410,260
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/410,259
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/410,259
; PRIOR FILING DATE: 2002-09-12
; NUMBER OF SEQ ID NOS: 5444
; SOFTWARE: PERL Program
; SEQ ID NO 5429
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; FEATURE:
; OTHER INFORMATION: Incyte ID No: 966243.PT7p
PCT-US03-28227-5429

Query Match 66.7%; Score 30; DB 1; Length 445;
Best Local Similarity 50.0%; Pred. No. 2e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0;

QY 1 LMAQALAPL 10
   :|:|:|:|
Db 266 IMVKEAISFL 275

RESULT 13
PCT-US03-28227-5431
; Sequence 5431, Application PC/TUS0328227
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; SCHMIDT, Jeanette P.;
; APPLICANT: WRIGHT, Rachel J.; BRUNS, Christopher M.;
; APPLICANT: MARJANOVIC, Mirjana M.; SHEN, Fan;
; APPLICANT: HARTSHORNE, Toinette A.; SUCHOROLSKI, Martin;
; APPLICANT: ALTUS, Christina M.; PITTS, Steven J.;
; APPLICANT: ELDER, Linda V.; MOONEY, Elizabeth M.;
; APPLICANT: DELEGEANE, Angelo M.; PANESAR, Iqbal S.;
; APPLICANT: BANVILLE, Steven C.; REDDY, Thirupathi P.;
; APPLICANT: STEVENS, Kristian A.; BLANCHARD, John L.;
; APPLICANT: PANZER, Scott R.; WANG, Xinhao;
; APPLICANT: AU, Alan P.; GERSTIN, Edward H., Jr.;
; APPLICANT: PERALTA, Careyna H.; ANDERSON, Scott E.;
; APPLICANT: RIOUX, Pierre; SHEN, Edward J.;
; APPLICANT: WU, Mingham C.; STUVE, Laura L.;
; APPLICANT: LAGACE, Robert E.; SPIRO, Peter A.;
; APPLICANT: STEWART, Elizabeth A.; WINGROVE, James A.;
; APPLICANT: VITT, Ursula A.; KIRTON, Edward;
; APPLICANT: XU, Yuming; KWONG, Mary;
; APPLICANT: POLICKY, Jennifer L.; HURWITZ, Bonnie L.;
; APPLICANT: MA, Yan; JACKSON, Jennifer L.;
; APPLICANT: GIETZEN, Darryl; PATURY, Srikanth;
; APPLICANT: SHI, Xiaobing; SUAREZ, Charlyn J.
; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PN-0100 PCT
; CURRENT APPLICATION NUMBER: US 60/410,260
; PRIOR FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: US 60/410,260
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/410,259
; PRIOR FILING DATE: 2002-09-12
; NUMBER OF SEQ ID NOS: 5444
; SOFTWARE: PERL Program
; SEQ ID NO 5431
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; FEATURE:
; OTHER INFORMATION: Incyte ID No: 966243.PT7p
PCT-US03-28227-5431

Query Match 66.7%; Score 30; DB 1; Length 445;
Best Local Similarity 50.0%; Pred. No. 2e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0;

QY 1 LMAQALAPL 10
   :|:|:|:|
Db 266 IMVKEAISFL 275

RESULT 12
PCT-US03-28227-5430
; Sequence 5430, Application PC/TUS0328227
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; SCHMIDT, Jeanette P.;
; APPLICANT: WRIGHT, Rachel J.; BRUNS, Christopher M.;
; APPLICANT: MARJANOVIC, Mirjana M.; SHEN, Fan;
; APPLICANT: HARTSHORNE, Toinette A.; SUCHOROLSKI, Martin;
; APPLICANT: ALTUS, Christina M.; PITTS, Steven J.;
; APPLICANT: ELDER, Linda V.; MOONEY, Elizabeth M.;
; APPLICANT: DELEGEANE, Angelo M.; PANESAR, Iqbal S.;
; APPLICANT: BANVILLE, Steven C.; REDDY, Thirupathi P.;
; APPLICANT: STEVENS, Kristian A.; BLANCHARD, John L.;
; APPLICANT: PANZER, Scott R.; WANG, Xinhao;
; APPLICANT: AU, Alan P.; GERSTIN, Edward H., Jr.;
; APPLICANT: PERALTA, Careyna H.; ANDERSON, Scott E.;
; APPLICANT: RIOUX, Pierre; SHEN, Edward J.;
; APPLICANT: WU, Mingham C.; STUVE, Laura L.;
; APPLICANT: LAGACE, Robert E.; SPIRO, Peter A.;
; APPLICANT: STEWART, Elizabeth A.; WINGROVE, James A.;
; APPLICANT: VITT, Ursula A.; KIRTON, Edward;

```

; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No: 966243.PT9p  
PCT-US03-28227-5431

Query Match 66.7%; Score 30; DB 1; Length 445;  
Best Local Similarity 50.0%; Pred. No. 2e+02;  
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMAQEALAF 10  
Db 266 IMVKRAISFL 275

## RESULT 14

PCT-US02-35618B-25  
; Sequence 25, Application PC/TUS0235618B  
; GENERAL INFORMATION:  
; APPLICANT: GEORGIOU, GEORGE  
; APPLICANT: DELISA, MATTHEW  
; TITLE OF INVENTION: ENGINEERING OF LEADER PEPTIDES FOR THE SECRETION OF  
; FILE REFERENCE: CLPR:019WO  
; CURRENT APPLICATION NUMBER: PCT/US02/35618B  
; CURRENT FILING DATE: 2002-11-05  
; PRIOR APPLICATION NUMBER: 10/289,135  
; PRIOR FILING DATE: 2002-11-05  
; PRIOR APPLICATION NUMBER: 60/337,452  
; PRIOR FILING DATE: 2001-11-05  
; NUMBER OF SEQ ID NOS: 134  
; SOFTWARE: PatentIn ver. 2.1  
; SEQ ID NO 25  
; LENGTH: 29  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
PCT-US02-35618B-25

Query Match 64.4%; Score 29; DB 1; Length 29;  
Best Local Similarity 70.0%; Pred. No. 45;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 LMAQEALAF 10  
Db 12 LTASSALAF 21

## RESULT 15

US-10-425-114A-38464  
; Sequence 38464, Application US/10425114A  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114A  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 38464  
; LENGTH: 135  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700351058\_FLI.pep  
US-10-425-114A-38464

Query Match 64.4%; Score 29; DB 6; Length 135;  
Best Local Similarity 85.7%; Pred. No. 1.3e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 QEALAF 10  
Db 24 EEALAF 30

Search completed: October 7, 2003, 13:55:55  
Job time : 3.16456 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:04:57 ; Search time 14.9241 Seconds  
(without alignments)  
95,721 Million cell updates/sec

Title: US-09-807-512-24  
Perfect score: 43  
Sequence: 1 FLMAQGAML 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues  
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_19Jun03:\*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	43	100.0	9	21 AAY70857	CAMEL10 immunogeni
2	43	100.0	30	23 AAU85114	Human NYN501b segm
3	43	100.0	58	20 AAY05966	Human cancer antig
4	43	100.0	58	21 AAY70863	Human tumour antig
5	43	100.0	58	23 AAU84819	Human NYN501b cons
6	43	100.0	109	21 AAY70854	Human CTL-recognis
7	43	100.0	3541	23 AAU85130	Human melanoma spe
8	32	74.4	181	22 ABG02039	Novel human diagn
9	32	74.4	406	21 AAB18987	Amino acid sequenc

10	32	74.4	1101	24 ABR47437	Breast cancer asso
11	32	74.4	1360	24 AAE32114	Human cytoskeleton
12	31	72.1	88	21 AAB43882	Human cancer assoc
13	31	72.1	88	22 AAG74103	Human colon cancer
14	31	72.1	222	22 ABB67760	Drosophila melanog
15	31	72.1	423	21 AAG30606	Arabidopsis thalia
16	31	72.1	423	21 AAG51411	Arabidopsis thalia
17	31	72.1	439	21 AAG30605	Arabidopsis thalia
18	31	72.1	439	21 AAG51410	Arabidopsis thalia
19	31	72.1	476	21 AAG30604	Arabidopsis thalia
20	31	72.1	476	21 AAG51409	Arabidopsis thalia
21	31	72.1	592	17 AAR98888	Nia protease of th
22	31	72.1	592	17 AAR98889	Nia protease of th
23	31	72.1	775	23 AAG19446	Novel human diagn
24	30	69.8	146	23 ABB89194	Human polypeptide
25	30	69.8	269	22 ABB67759	Drosophila melanog
26	30	69.8	302	22 AAB94000	Human protein sequ
27	30	69.8	402	22 AAB85362	Human phosphatase
28	30	69.8	411	23 ABB69581	Human polypeptide
29	30	69.8	434	22 AAM25227	Human protein sequ
30	30	69.8	471	19 AAW61216	Streptococcus pneu
31	30	69.8	471	23 AAB54635	S. pneumoniae SP09
32	30	69.8	490	21 AAY81745	Streptococcus pneu
33	30	69.8	506	24 ABU02034	S. pneumoniae type
34	30	69.8	537	22 ABB64617	Drosophila melanog
35	30	69.8	773	22 ABB62444	Drosophila melanog
36	30	69.8	1011	23 ABB73711	Candida albicans e
37	30	69.8	1068	21 AAY96750	A. terreus ORF8 HM
38	29	67.4	59	23 ABB32917	Human ORF1890 prot
39	29	67.4	62	21 AAB16374	Eucalyptus grandis
40	29	67.4	78	24 ABB76304	Human GENSET prote
41	29	67.4	103	23 ABB31533	Human ORF506 prote
42	29	67.4	103	23 ABB32558	Human glycoprotein
43	29	67.4	140	20 AAW66331	Kidney injury asso
44	29	67.4	241	23 ABB09512	Human TSPAN-1-like
45	29	67.4	264	23 ABB63098	FLO11 gene express

ALIGNMENTS

RESULT 1  
AAY70857  
ID AAY70857 standard; peptide; 9 AA.

AC AAY70857;

XX  
DT 31-JUL-2000 (first entry)

DE CAMEL10 immunogenic peptide of human CAMEL protein.

XX  
KW CAMEL: CTL-recognised Antigen on MELANOMA; cytotoxic T lymphocyte; CTL;  
KW tumour-associated antigen; LAGE-1; NY-ESO-1; anticancer; melanoma;  
KW human; cancer; immunotherapy; immunogenic peptide; immune response.

XX Homo sapiens.

PN WO200023584-A1.

PD 27-APR-2000.

PF 15-OCT-1999; 99WO-EP07832.

PR 16-OCT-1998; 98EP-0119583.

XX (BOEH ) BOEHRINGER INGELHEIM INT GMBH.  
PA (UYHO-) UNIV HOSPITAL LEIDEN.

PI Schrier PI, Aarnoudse CA, Heider K, Klade C;

XX WPI; 2000-339685/29.

PT Tumor-associated antigen useful for cancer immunotherapy is encoded by

PT the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA -  
XX Claim 6; Page 34; 73pp; English.  
PS  
CC The present sequence is an immunogenic peptide CAMEL 10, of the human  
XX tumour-associated antigen CAMEL (Cytotoxic T lymphocytes (CTL)-recognised  
CC Antigen on MELANOMA). This peptide has the potential to bind to HLA-A2  
CC and corresponds to residues 10-18 of the CAMEL protein. The CAMEL protein  
CC is encoded by the LAGE-1 gene, a tumour-specific antigen. It is different  
CC from the LAGE-1 protein, since it is translated from a different open  
CC reading frame (ORF-1). It shows strong homology with NY-ESO-1, a melanoma  
CC specific tumour antigen. The tumour-associated antigen displayed on  
CC melanoma cells is recognised by cytotoxic T lymphocytes. CAMEL is  
CC expressed in tumour cell lines, tumour tissues (e.g. breast and lung) and  
CC in restricted number of healthy tissues. This sequence has anticancer  
CC activity. CAMEL tumour antigen and immunogenic peptides derived from it  
CC are useful for cancer immunotherapy. They have the potential to induce an  
CC immune response, by eliciting a CTL response. The DNA molecule is used to  
CC construct recombinant or fusion proteins.  
XX  
XX  
SQ Sequence 9 AA;  
Query Match 100.0%; Score 43; DB 21; Length 9;  
Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 FLMAQGAML 9  
Db 1 FLMAQGAML 9  
IIIIIIII  
RESULT 2  
AAU85114  
ID AAU85114 standard; Peptide; 30 AA.  
XX  
XX AAU85114;  
AC  
XX  
XX 08-MAY-2002 (first entry)  
DT  
XX  
XX Human NYNSolb segment 1.  
DE  
XX  
XX Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;  
KW viral infection; human immunodeficiency virus; melanoma;  
KW bacterial infection; Salmonella; Legionella; parasitic infection;  
KW Trypanosoma; Toxoplasma; Giardia.  
XX  
XX Homo sapiens.  
OS  
XX WO200190197-A1.  
PN  
XX 29-NOV-2001.  
PD  
XX  
XX 25-MAY-2001; 2001WO-AU00622.  
PF  
XX  
XX 26-MAY-2000; 2000AU-0007761.  
PR  
XX  
XX (AUSU ) UNIV AUSTRALIAN NAT.  
PA  
XX  
XX Thomson SA, Ramshaw IA;  
PI  
XX  
XX WPI; 2002-147575/19.  
DR  
XX N-PSDB; ABK36934.  
XX  
XX New synthetic polypeptides having several different segments of at  
PT least one parent polypeptide linked together differently compared to  
PT the linkage in the parent polypeptide, for inducing immune response  
PT against a pathogen or cancer  
XX  
XX Example 3; Fig 27; 364pp; English.  
PS  
XX  
XX The invention relates to a new synthetic polypeptide (I) comprising  
CC several different segments of at least one parent polypeptide linked  
CC together in a different relationship relative to their linkage in the

CC parent polypeptide to impede, abrogate or otherwise alter at least one  
CC function associated with the parent polypeptide and for inducing an  
CC immune response against a pathogen or cancer. Also included are a  
CC synthetic polynucleotide encoding and a computer system for  
CC designing the synthetic polypeptides. The synthetic polypeptides and  
CC polynucleotides are referred to as a Savine. The synthetic polypeptide is  
CC useful for modulating immune responses preferably directed against a  
CC pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, bone  
CC colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone  
CC liver, oesophagus, brain, testicle, uterus), as potentiating agents.  
CC Compositions comprising the polypeptide may be used in the treatment or  
CC prophylaxis against viral (such as infections caused by HIV (human  
CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis  
CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial  
CC (e.g., infections caused by Neisseria, Meningococcus, Haemophilus,  
CC Salmonella, Streptococcus, Legionella and Mycobacterium or parasitic  
CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,  
CC Trypanosoma, Toxoplasma and Giardia) infections. The present  
CC sequence is a peptide derived from a parent protein used to  
CC construct a savine of the invention.  
XX  
XX  
SQ Sequence 30 AA;  
Query Match 100.0%; Score 43; DB 23; Length 30;  
Best Local Similarity 100.0%; Pred. No. 0.029;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 FLMAQGAML 9  
Db 12 FLMAQGAML 20  
IIIIIIII  
RESULT 3  
AAY05966  
ID AAY05966 standard; Protein; 58 AA.  
XX  
XX AAY05966;  
AC  
XX  
XX 16-AUG-1999 (first entry)  
DT  
XX  
XX Human cancer antigen NY ESO-1/CAG-3 ORF2 protein.  
DE  
XX  
XX NY ESO-1/CAG-3 gene; CAG-3 gene; cancer antigen; human; leukaemia;  
KW non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
KW vaccine; ORF2.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO9918206-A2.  
PN  
XX  
XX 15-APR-1999.  
PD  
XX  
XX 21-SEP-1998; 98WO-US19609.  
PF  
XX  
XX 08-OCT-1997; 97US-0061428.  
PR  
XX  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
PA  
XX  
XX Rosenberg SA, Wang RF;  
PI  
XX  
XX WPI; 1999-277270/23.  
DR  
XX N-PSDB; AAX58599.  
XX  
XX Cancer antigen NY ESO1/CAG-3  
PT  
XX  
XX Claim 5; Fig 3A; 88pp; English.  
PS  
XX  
XX The present sequence represents the ORF2 protein encoded by  
CC open reading frame 2 of the human ESO-1/CAG-3 (or CAG-3) gene.

CC CAG-3 is a new and potent tumour antigen capable of eliciting an  
 CC antigen specific immune response by T cells. Cancer peptides  
 CC comprising ORF2, ORF1 (see AAY05965), portions of these peptides and  
 CC their variants (see AAY05967-87), are useful as cancer vaccines that  
 CC protect the recipient from development of cancer. The invention  
 CC provides: vectors and host cells (also useful as vaccines); a  
 CC method of diagnosis of cancer or precancer; a transgenic animal;  
 CC antisense oligonucleotides that inhibit expression of the cancer  
 CC peptide or tumour antigen; antibodies reacting with the CAG-3  
 CC cancer peptide, useful in diagnostic and detection assays; and  
 CC methods for preventing or inhibiting cancer by administering a  
 CC cancer peptide, with or without an HLA molecule. The cancer  
 CC peptides form part of, or are derived from, cancers such as  
 CC primary or metastatic melanoma, thymoma, lymphoma, sarcoma, lung  
 CC cancer, liver cancer, leukaemia, uterine cancer, cervical cancer,  
 CC bladder cancer, kidney cancer and adenocarcinomas such as breast,  
 CC prostate, ovarian, pancreatic and thyroid cancers. Melanoma is  
 CC treated by inducing cancer-specific T cells in vitro for subsequent  
 CC return to a patient.

XX  
 SQ Sequence 58 AA;

Query Match 100.0%; Score 43; DB 20; Length 58;  
 Best Local Similarity 100.0%; Pred. No. 0.059;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9  
 |||||  
 Db 10 FLMAQGAML 18

RESULT 4  
 AAY70863  
 ID AAY70863 standard; Protein; 58 AA.  
 AC AAY70863;  
 XX  
 DT 31-JUL-2000 (first entry)  
 XX  
 DE Human tumour antigen, NY-ESO-1 short variant protein.  
 XX  
 KW NY-ESO-1; CAMEL; CTL-recognised Antigen on MELANOMA; human; cancer; CTL;  
 KW cytotoxic T lymphocyte; tumour-associated antigen; LAGE-1; anticancer;  
 KW melanoma; immunotherapy; immune response; variant.  
 OS Homo sapiens.  
 XX  
 XX WO200023584-A1.  
 XX  
 PD 27-APR-2000.  
 XX  
 PF 15-OCT-1999; 99WO-EP07832.  
 XX  
 PR 16-OCT-1998; 98EP-0119583.  
 XX  
 PA (BOEH ) BOEHRINGER INGELHEIM INT GMBH.  
 XX (UYHO-) UNIV HOSPITAL LEIDEN.  
 XX  
 PI Schrier PI, Aarnoudse CA, Heider K, Klade C;  
 XX  
 DR WPI; 2000-339685/29.  
 DR N-PSDB; AAD00152.  
 XX  
 PT Tumor-associated antigen useful for cancer immunotherapy is encoded by  
 PT the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA -  
 XX  
 XX Example 3; Page 65; 73pp; English.  
 CC The present sequence is the human NY-ESO-1 short variant protein, a  
 CC tumour antigen, identified by screening an esophagus carcinoma cDNA  
 CC library. This protein is derived from open reading frame (ORF)-2,  
 CC that differs from the CAMEL protein (cytotoxic T lymphocytes (CTL)-  
 CC recognised Antigen on MELANOMA), a tumour-associated antigen, by only its

CC last 5 amino acids. It contains epitopes of tumour specific T-cells.  
 CC NY-ESO-1 is expressed in different tumour types, but not in healthy  
 CC tissues except in testis. The tumour-associated antigen displayed on  
 CC melanoma cells is recognised by cytotoxic T lymphocytes. This sequence  
 CC has anticancer activity. CAMEL tumour antigen and immunogenic peptides  
 CC derived from it are useful for cancer immunotherapy. They have the  
 CC potential to induce an immune response, by eliciting a CTL response.  
 CC The DNA molecule is used for the construction of recombinant or fusion  
 CC proteins.

XX  
 SQ Sequence 58 AA;

Query Match 100.0%; Score 43; DB 21; Length 58;  
 Best Local Similarity 100.0%; Pred. No. 0.059;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9  
 |||||  
 Db 10 FLMAQGAML 18

RESULT 5  
 AAU84819  
 ID AAU84819 standard; Protein; 58 AA.  
 XX  
 AC AAU84819;  
 XX  
 DT 08-MAY-2002 (first entry)  
 XX  
 DE Human NYNS01b consensus sequence.  
 XX  
 KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;  
 KW viral infection; human immunodeficiency virus; melanoma;  
 KW bacterial infection; Salmonella; Legionella; parasitic infection;  
 KW Trypanosoma; Toxoplasma; Giardia.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200190197-A1.  
 XX  
 PD 29-NOV-2001.  
 XX  
 PF 25-MAY-2001; 2001WO-AU00622.  
 XX  
 PR 26-MAY-2000; 2000AU-0007761.  
 XX  
 PA (AUSU ) UNIV AUSTRALIAN NAT.  
 XX  
 PI Thomson SA, Ramshaw IA;  
 XX  
 DR WPI; 2002-147575/19.  
 XX  
 PT New synthetic polypeptides having several different segments of at  
 PT least one parent polypeptide linked together differently compared to  
 PT the linkage in the parent polypeptide, for inducing immune response  
 PT against a pathogen or cancer  
 XX  
 XX Example 3; Fig 27; 364pp; English.  
 CC The invention relates to a new synthetic polypeptide (I) comprising  
 CC several different segments of at least one parent polypeptide linked  
 CC together in a different relationship relative to their linkage in the  
 CC parent polypeptide to impede, abrogate or otherwise alter at least one  
 CC function associated with the parent polypeptide and for inducing an  
 CC immune response against a pathogen or cancer. Also included are a  
 CC synthetic polynucleotide encoding and a computer system for  
 CC designing the synthetic polypeptides. The synthetic polypeptides and  
 CC polynucleotides are referred to as a Savine. The synthetic polypeptide is  
 CC useful for modulating immune responses preferably directed against a  
 CC pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix,  
 CC colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone  
 CC liver, oesophagus, brain, testicle, uterus), as potentiating agents.  
 CC Compositions comprising the polypeptide may be used in the treatment or

CC prophylaxis against viral (such as infections caused by HIV (human  
CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis  
CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial  
CC (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,  
CC Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic  
CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,  
CC Trypanosoma, Toxoplasma and Giardia) infections. The present  
CC sequence is a consensus sequence for a parent protein used to design a  
CC savine of the invention.

XX  
XX  
SQ Sequence 58 AA;  
Query Match 100.0%; Score 43; DB 23; Length 58;  
Best Local Similarity 100.0%; Pred. No. 0.059; 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;

QY 1 FLMAQGAML 9  
DB 10 FLMAQGAML 18  
|||||

RESULT 6  
AAV70854  
ID AAV70854 standard; Protein: 109 AA.  
XX AC AAV70854;  
XX  
XX 31-JUL-2000 (first entry)  
XX Human CTL-recognised Antigen on MELanoma (CAMEL) protein.  
XX CAMEL; CTL-recognised Antigen on MELanoma; cytotoxic T lymphocyte; CTL;  
KW tumour-associated antigen; LAGE-1; NY-ESO-1; anticancer; melanoma;  
KW cancer; immunotherapy; immunogenic peptide; Immune response.  
XX  
OS Homo sapiens.

XX  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..11 /label= CTL\_epitope  
FT /note= "Immunogenic peptide with ability to elicit a  
FT CTL response"  
FT Peptide 2..10 /label= Immunogenic\_peptide  
FT /note= "Specific for HLA-A3"  
FT Peptide 2..11 /label= CTL\_epitope  
FT /note= "Immunogenic peptide with ability to elicit a  
FT CTL response"  
FT Peptide 10..18 /label= CAMEL\_10  
FT /note= "Specific for HLA-A2"  
FT Peptide 16..25 /label= CAMEL\_16  
FT /note= "Specific for HLA-A2"  
FT Peptide 17..25 /label= CAMEL\_17  
FT /note= "Specific for HLA-A2"  
FT Peptide 51..59 /label= Immunogenic\_peptide  
FT /note= "Specific for HLA-A3101"  
FT Peptide 101..109 /label= Immunogenic\_peptide  
XX  
XX WO200023584-A1.  
XX  
XX 27-APR-2000.  
XX  
XX 15-OCT-1999; 99WO-EF07832.  
XX  
XX 16-OCT-1998; 98EP-0119583.  
XX  
XX (BOEH ) BOEHRINGER INGELHEIM INT GMBH.  
PA

PA (UYHO-) UNIV HOSPITAL LEIDEN.  
XX Schrier PI, Aarnoudse CA, Heider K, Klade C;  
XX WPI; 2000-339685/29.  
DR N-PSDB; AAD00149.  
XX  
XX Tumor-associated antigen useful for cancer immunotherapy is encoded by  
PT the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA -  
XX  
XX Claim 1; Page 55; 73pp; English.  
XX  
XX The present protein sequence is the human tumour-associated antigen CAMEL  
CC (Cytotoxic T lymphocytes (CTL)-recognised Antigen on MELanoma). CAMEL  
CC protein is encoded by the LAGE-1 gene, a tumour-specific antigen. It is  
CC different from the LAGE-1 protein, since it is translated from a  
CC different open reading frame (ORF-1). It shows strong homology with  
CC NY-ESO-1, a melanoma specific tumour antigen. The tumour-associated  
CC antigen displayed on melanoma cells is recognised by cytotoxic T  
CC lymphocytes. CAMEL is expressed in tumour cell lines, tumour tissues  
CC (e.g. breast and lung) and in restricted number of healthy tissues. This  
CC sequence has anticancer activity. CAMEL tumour antigen and immunogenic  
CC peptides derived from it are useful for cancer immunotherapy. They have  
CC the potential to induce an immune response, by eliciting a CTL response.  
CC The DNA molecule is used to construct recombinant or fusion proteins.  
XX  
XX SQ Sequence 109 AA;  
Query Match 100.0%; Score 43; DB 21; Length 109;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9  
DB 10 FLMAQGAML 18  
|||||

RESULT 7  
AAU85130  
ID AAU85130 standard; Protein: 3541 AA.  
XX AC AAU85130;  
XX  
XX 08-MAY-2002 (first entry)  
XX Human melanoma specific savine.  
XX  
XX Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;  
KW viral infection; human immunodeficiency virus; melanoma;  
KW bacterial infection; Salmonella; Legionella; parasitic infection;  
KW Trypanosoma; Toxoplasma; Giardia.  
XX  
XX Homo sapiens.  
OS Synthetic.  
XX  
XX WO200190197-A1.  
XX  
XX 29-NOV-2001.  
XX  
XX 25-MAY-2001; 2001WO-AU00622.  
XX  
XX 26-MAY-2000; 2000AU-0007761.  
XX  
XX (AUSU ) UNIV AUSTRALIAN NAT.  
XX  
XX Thomson SA, Ramshaw IA;  
XX  
XX WPI; 2002-147575/19.  
DR N-PSDB; ABK36950.  
XX  
XX New synthetic polypeptides having several different segments of at  
PT least one parent polypeptide linked together differently compared to  
PT the linkage in the parent polypeptide, for inducing immune response



PT against a pathogen or cancer -

XX Example 3; Fig 27; 364pp; English.

XX The invention relates to a new synthetic polypeptide (I) comprising

CC several different segments of at least one parent polypeptide linked

CC together in a different relationship relative to their linkage in the

CC parent polypeptide to impede, abrogate or otherwise alter at least one

CC function associated with the parent polypeptide and for inducing an

CC immune response against a pathogen or cancer. Also included are a

CC synthetic polynucleotide encoding and a computer system for

CC designing the synthetic polypeptides. The synthetic polypeptides and

CC polynucleotides are referred to as a Savine. The synthetic polypeptide is

CC useful for modulating immune responses preferably directed against a

CC pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix,

CC colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone

CC liver, oesophagus, brain, testicle, uterus), as potentiating agents.

CC Compositions comprising the polypeptide may be used in the treatment or

CC prophylaxis against viral (such as infections caused by HIV (human

CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis

CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial

CC (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,

CC Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic

CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,

CC Trypanosoma, Toxoplasma and Giardia) infections. The present

CC sequence is a savine protein of the invention.

XX

SQ Sequence 3541 AA;

Query Match 100.0%; Score 43; DB 23; Length 3541;

Best Local Similarity 100.0%; Pred. No. 4.9;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9

Db 1763 FLMAQGAML 1771

|||||||

RESULT 8

ABG02039

ID ABG02039 standard; Protein; 181 AA.

XX

AC ABG02039;

XX

DT 13-FEB-2002 (first entry)

XX

DE Novel human diagnostic protein #2030.

XX

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

KW

XX Homo sapiens.

OS

XX WO200175067-A2.

PN

XX

PD 11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US08631.

XX

XX 31-MAR-2000; 2000US-0540217.

PR

XX 23-AUG-2000; 2000US-0649167.

XX

PA (HYSE-) HYSEQ INC.

XX

XX Drmanac RT, Liu C, Tang YT;

PI

XX WPI; 2001-639362/73.

DR

DR N-PSDB; AA566226.

XX

PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity -

XX Claim 20; SEQ ID No 32398; 103pp; English.

PS

XX The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques

CC to restore normal activity of (II) or to treat disease states involving

CC (II). (II) is useful for generating antibodies against it, detecting or

CC quantitating a polypeptide in tissue, as molecular weight markers and as

CC a food supplement. (II) and its binding partners are useful in medical

CC imaging of sites expressing (II). (I) and (II) are useful for treating

CC disorders involving aberrant protein expression or biological activity.

CC The polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG00010-ABG30377 represent novel human

CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX

SQ Sequence 181 AA;

Query Match 74.4%; Score 32; DB 22; Length 181;

Best Local Similarity 66.7%; Pred. No. 49;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9

Db 140 FIMVNGAML 148

|||||

RESULT 9

AAB18987

ID AAB18987 standard; Protein; 406 AA.

XX

AC AAB18987;

XX

DT 08-FEB-2001 (first entry)

XX

XX Amino acid sequence of a human transmembrane protein.

XX

XX Human; transmembrane protein; cell proliferation disorder; myeloma;

KW reproductive disorder; smooth muscle disorder; neurological disorder;

KW arteriosclerosis; leukaemia; acquired immunodeficiency syndrome; AIDS;

KW allergy; ovulatory defect; angina; hypertension; stroke; epilepsy;

KW Alzheimer's disease; Tourette's disorder.

XX

XX Homo sapiens.

OS

XX

FH Key Location/Qualifiers

FT Modified-site 4

FT /note= "potential phosphorylation site"

FT Modified-site 31

FT /note= "potential glycosylation site"

FT Modified-site 59

FT /note= "potential phosphorylation site"

FT Modified-site 87

FT /note= "potential phosphorylation site"

FT Modified-site 94

FT /note= "potential phosphorylation site"

FT Modified-site 99

FT /note= "potential phosphorylation site"

FT Modified-site 112

FT /note= "potential glycosylation site"

FT Modified-site 113

FT /note= "potential phosphorylation site"

FT Modified-site 125

FT /note= "potential phosphorylation site"

FT

```

FT Modified-site 136
FT /note= "potential phosphorylation site"
FT Modified-site 281
FT /note= "potential phosphorylation site"
FT Modified-site 365
FT /note= "potential phosphorylation site"
XX
XX WO200056891-A2.
XX
XX 28-SEP-2000.
XX
XX 22-MAR-2000; 2000WO-US07817.
XX
XX 22-MAR-1999; 99US-0125537.
XX 16-JUN-1999; 99US-0139565.
XX
XX (INCY-) INCYTE PHARM INC.
XX
XX Yue H, Lal P, Tang YT, Hillman JL, Reddy R, Bandman O, Baughn MR;
XX Lu DAM, Azimzai Y, Yang J;
XX
XX WPI; 2000-579485/54.
XX N-PSDB; AAA96500.
XX
XX New human transmembrane proteins are used to treat a disease or
XX condition associated with decreased expression of functional HTMP e.g.
XX Tourette's disorder, angina and leukaemia -
XX
XX Claim 1; Page 104-105; 130pp; English.
XX
XX The present sequence represents a human transmembrane proteins (HTMP).
XX Agonists and antagonists of the protein are used to treat a disease
XX or condition associated with overexpression of the protein. Diseases
XX and conditions which can be treated include cell proliferative,
XX immunological, reproductive, smooth muscle and neurological disorders
XX e.g. arteriosclerosis, myeloma, leukaemia, acquired immunodeficiency
XX syndrome (AIDS), allergies, ovulatory defects, angina, hypertension,
XX stroke, Alzheimer's disease, epilepsy and Tourette's disorder. The
XX polynucleotides may be used to detect and quantify gene expression in
XX biopsied tissues where protein expression may be correlated with disease
XX e.g. to determine absence, presence or excess expression of HTMP or to
XX monitor regulation of HTMP expression during therapeutic intervention.
XX
XX Sequence 406 AA;
XX
XX Query Match 74.4%; Score 32; DB 21; Length 406;
XX Best Local Similarity 66.7%; Pred No. 1.2e+02;
XX Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 FLMAQGAML 9
XX I:| | | | |
XX Db 302 FIMVNGAML 310
XX
XX RESULT 10
XX ABR47437
XX ID ABR47437 standard; Protein; 1101 AA.
XX
XX AC ABR47437;
XX
XX DT 12-JUN-2003 (first entry)
XX
XX DE Breast cancer associated protein sequence SEQ ID NO:106.
XX
XX KW Human; breast cancer; cytostatic; gene therapy.
XX
XX OS Homo sapiens.
XX
XX PN WO2003004989-A2.
XX
XX PD 16-JAN-2003.
XX
XX PF 21-JUN-2002; 2002WO-US19669.

```

```

XX 21-JUN-2001; 2001US-299887P.
XX 27-JUN-2001; 2001US-301572P.
XX 18-JUL-2001; 2001US-306501P.
XX 25-SEP-2001; 2001US-325002P.
XX 05-MAR-2002; 2002US-362585P.
XX 14-MAY-2002; 2002US-380391P.
XX
XX (MILL-) MILLENIUM PHARM INC.
XX
XX Lillie J, Gannavarapu M, Glatt K, Hoersch S, Kamatkar S, Mertens M;
XX Monahan JE, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE, Bast RC;
XX Hortobagyi GN, Pusztai L, Meric F, Sahin A, Mills GB;
XX
XX WPI; 2003-210381/20.
XX N-PSDB; ACC50129.
XX
XX Breast cancer diagnosis or treatment by comparing the level of
XX expression of a marker in a patient sample with that in the control
XX non-breast cancer sample -
XX
XX Claim 1; SEQ ID 106; 128pp; English.
XX
XX The present invention describes a method for assessing whether a patient
XX is afflicted with breast cancer. The method comprises comparing the level
XX of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and
XX ABR47386 to ABR47632) in a patient sample and the normal level of
XX expression of the marker in a control non-breast cancer sample, where a
XX significant increase in the level of expression of the marker in the
XX patient sample and the normal level is an indication that the patient is
XX afflicted with breast cancer. The breast cancer associated sequences
XX from the present invention have cytostatic activities and can be used in
XX gene therapy. The method is useful for diagnosing and treating breast
XX cancer.
XX N.B. The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 1101 AA;
XX
XX Query Match 74.4%; Score 32; DB 24; Length 1101;
XX Best Local Similarity 66.7%; Pred. No. 3.4e+02;
XX Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 FLMAQGAML 9
XX I:| | | | |
XX Db 219 FLAQGASI 227
XX
XX RESULT 11
XX AAE32114
XX ID AAE32114 standard; Protein; 1360 AA.
XX
XX AC AAE32114;
XX
XX DT 24-MAR-2003 (first entry)
XX
XX DE Human cytoskeleton-associated protein, CSAP-12.
XX
XX KW Human; cytoskeleton-associated protein; CSAP-12; atherosclerosis;
XX cancer; gene therapy.
XX
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX Peptide 1..22
XX /label= Signal_peptide
XX Protein 23..1360
XX /note= "Human mature CSAP-12"
XX Domain 289..313
XX /note= "Transmembrane domain"
XX Domain 360..376
XX /note= "Transmembrane domain"
XX

```

```

FT Domain 614..647
FT /note= "TPR domain"
FT Domain 661..694
FT /note= "TPR domain"
FT Domain 695..728
FT /note= "TPR domain"
XX WO200279404-A2.
XX 10-OCT-2002.
XX 25-MAR-2002; 2002WO-US09288.
XX 29-MAR-2001; 2001US-280508P.
PR 03-APR-2001; 2001US-281323P.
PR 13-APR-2001; 2001US-283769P.
PR 04-MAY-2001; 2001US-288609P.
PR 10-MAY-2001; 2001US-290518P.
PR 18-MAY-2001; 2001US-291870P.
PR 29-MAY-2001; 2001US-294451P.
XX (INCY-) INCYTE GENOMICS INC.
XX Hafalia AJA, Tang TY, Yue H, Khan FA, Ison CH, Baughn MR;
PI Warren BA, Duggan BM, Thangavelu K, Honchell CD, Azimzai Y;
PI Elliott VS, Burford N, Ding L, Yue H, Becha S, Emerling BM;
PI Richardson TW, Lee SY, Bandman O, Lal PG, Lee S, Gietzen KJ;
PI Walla NK, Griffin JA, Lee EA, Swarnakar A, Ring HZ, Jones KA;
XX WPI: 2003-092894/08.
DR N-PSDB; AAD49601.
XX
PT New human cytoskeleton-associated proteins, useful for preparing a
PT composition for diagnosing or treating a disease or condition
PT associated with decreased expression or overexpression of functional
PT CSAP e.g., cancer
XX
PS Claim 1; Page 171-174; 233pp; English.
XX
CC The invention relates to new human cytoskeleton-associated protein
CC (CSAP) and its polynucleotide. The polypeptide is useful for preparing
CC a composition for diagnosing or treating a disease or condition
CC associated with decreased expression or overexpression of functional
CC CSAP e.g. atherosclerosis or cancer. The present sequence is human
CC CSAP-12 protein. The invention is useful in gene therapy.
XX
SQ Sequence 1360 AA;
Query Match 74.4%; Score 32; DB 24; Length 1360;
Best Local Similarity 66.7%; Pred. No. 4.3e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
OY 1 FLMAQGAML 9
Db 488 FLAQQGASI 496
RESULT 12
AAB43882
ID AAB43882 standard; Protein; 88 AA.
XX
AC AAB43882;
XX
XX 08-FEB-2001 (first entry)
DE Human cancer associated protein sequence SEQ ID NO:1327.
XX
KW Human; cancer associated gene; cancer antigen; detection; cancer;
KW diagnosis; cytostatic; proliferative; vulnerary; immunomodulator;
KW antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
KW antinflammatory; antithyroid; antiallergic; antibacterial; cardiac;
KW dermatological; neuroprotective; thrombolytic; coagulant; neutropic;
KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;

```

---

```

KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW allergic reaction; graft versus host disease; organ rejection;
KW haemostatic; thrombolytic; cardiovascular disorder; infection;
KW neurological disease; drug screening.
XX
XX Homo sapiens.
XX WO200055350-A1.
XX 21-SEP-2000.
XX
XX 08-MAR-2000; 2000WO-US05882.
XX 12-MAR-1999; 99US-0124270.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX WPI: 2000-587533/55.
XX N-PSDB; AAC78091.
XX
XX Novel isolated nucleic acids comprising sequences encoding peptides
XX useful for treating or diagnosing e.g. cancer
XX
XX Claim 11; Page 1978-1979; 2352pp; English.
XX
XX AAC77607 to AAC78448 encode the human cancer associated proteins given
XX in AAB43398 to AAB44239. The proteins can have activities based on the
XX tissues and cells the genes are expressed in. Example of activities
XX include: cytostatic; proliferative; vulnerary; immunomodulator;
XX antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
XX antinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
XX dermatological; neuroprotective; cardiac; thrombolytic; coagulant;
XX neutropic; vasotropic; antipsoriatic and antiangiogenic. The
XX polynucleotides and polypeptides can be used for preventing, treating or
XX ameliorating medical conditions and diagnosing pathological conditions.
XX Polynucleotides, polypeptides, antibodies, agonists and antagonists from
XX the present invention may be used to treat immune disorders by activating
XX or inhibiting the proliferation, differentiation or mobilisation of
XX immune cells, to treat disorders of haematopoietic cells, autoimmune
XX disorders, allergic reactions, graft versus host disease and organ
XX rejection, modulate haemostatic or thrombolytic activity, modulate
XX inflammation, cancers, cardiovascular disorders, neurological disease and
XX bacterial or viral infections. The peptides, nucleotides, antibodies,
XX agonists and antagonists may be also be used in drug screens AAC78449 to
XX AAC78457 and AAB44240 represent sequences used in the exemplification of
XX the present invention.
XX
XX Sequence 88 AA;
Query Match 72.1%; Score 31; DB 21; Length 88;
Best Local Similarity 66.7%; Pred. No. 38;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
OY 1 FLMAQGAML 9
Db 17 FLAALGAMI 25
RESULT 13
AAG74103
ID AAG74103 standard; Protein; 88 AA.
XX
AC AAG74103;
XX
XX 03-SEP-2001 (first entry)
DE Human colon cancer antigen protein SEQ ID NO:4867.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma.
XX

```

OS Homo sapiens.  
XX W0200122920-A2.  
PN  
XX 05-APR-2001.  
PD  
XX 28-SEP-2000; 2000WO-US26524.  
XX  
XX 29-SEP-1999; 99US-0157137.  
XX 03-NOV-1999; 99US-0163280.  
PR  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Ruben SM, Barash SC, Birse CE, Rosen CA;  
PI  
XX WPI: 2001-235357/24.  
XX N-PSDB; AAH33534.  
DR  
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,  
PT useful for preventing, diagnosing and/or treating colorectal cancers -  
PT  
XX Claim 11; Page 6649; 9803pp; English.  
PS  
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon  
CC cancer-associated nucleic acid molecules (N) and proteins (P), where  
CC the proteins are collectively known as colon cancer antigens. The colon  
CC cancer antigens have cytostatic activity and can be used in gene  
CC therapy and vaccine production. N and P may be used in the prevention,  
CC diagnosis and treatment of diseases associated with inappropriate P  
CC expression. For example, N and P may be used to treat disorders  
CC associated with decreased expression by rectifying mutations or deletions  
CC in a patient's genome that affect the activity of P by expressing  
CC inactive proteins or to supplement the patients own production of P.  
CC Additionally, N may be used to produce the colon cancer-associated Ps,  
CC by inserting the nucleic acids into a host cell and culturing the cell  
CC to express the proteins. N and P can be used in the prevention, diagnosis  
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204  
CC and AAH77789 represent sequences used in the exemplification of the  
CC present invention.  
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were  
CC missing at time of publication, meaning no sequences are present for  
CC SEQ ID NO:1027 to 1052, 7921 and 7922.  
XX  
XX Sequence 88 AA;  
SQ  
Query Match 72.1%; Score 31; DB 22; Length 88;  
Best Local Similarity 66.7%; Pred. No. 38;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 1 FLMAQGAML 9  
DB 17 FLALGAMI 25  
RESULT 14  
ABB67760  
ID ABB67760 standard; Protein; 222 AA.  
XX  
XX ABB67760;  
AC  
XX 26-MAR-2002 (first entry)  
DT  
XX Drosophila melanogaster polypeptide SEQ ID NO 30072.  
DE  
XX Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
KW  
XX Drosophila melanogaster.  
OS  
XX W0200171042-A2.  
XX  
XX 27-SEP-2001.  
XX

PF 23-MAR-2001; 2001WO-US09231.  
XX  
XX 23-MAR-2000; 2000US-191637P.  
PR  
XX 11-JUL-2000; 2000US-0614150.  
PR  
XX (PEKE ) PE CORP NY.  
PA  
XX Venter JC, Adams M, Li PWD, Myers EW;  
PI  
XX WPI: 2001-656860/75.  
DR  
XX N-PSDB; ABL11863.  
DR  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
PT  
XX Disclosure; SEQ ID NO 30072; 21pp + Sequence Listing; English.  
PS  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins  
CC (ABB57737-ABB72072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 222 AA;  
SQ  
Query Match 72.1%; Score 31; DB 22; Length 222;  
Best Local Similarity 75.0%; Pred. No. 1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 FLMAQGAM 8  
DB 59 FLMAEGMM 66  
RESULT 15  
AAG30606  
ID AAG30606 standard; Protein; 423 AA.  
XX  
XX AAG30606;  
AC  
XX 17-OCT-2000 (first entry)  
DT  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 36620.  
DE  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
XX Arabidopsis thaliana.  
OS  
XX EP1033405-A2.  
PN  
XX 06-SEP-2000.  
PD  
XX 25-FEB-2000; 2000EP-0301439.  
PF  
XX 25-FEB-1999; 99US-0121825.  
XX 05-MAR-1999; 99US-0123180.  
PR  
XX 09-MAR-1999; 99US-0123548.  
PR  
XX 23-MAR-1999; 99US-0125788.  
PR  
XX 25-MAR-1999; 99US-0126264.  
PR  
XX 29-MAR-1999; 99US-0126785.  
PR  
XX 01-APR-1999; 99US-0127462.  
PR  
XX 06-APR-1999; 99US-0128234.  
PR  
XX 08-APR-1999; 99US-0128714.  
PR  
XX 16-APR-1999; 99US-0129845.  
PR

PR	19-APR-1999;	990S-0130077.	PR	19-JUL-1999;	990S-01443335.
PR	21-APR-1999;	990S-0130449.	PR	20-JUL-1999;	990S-0144352.
PR	23-APR-1999;	990S-0130510.	PR	20-JUL-1999;	990S-01446332.
PR	28-APR-1999;	990S-0130891.	PR	21-JUL-1999;	990S-0144884.
PR	28-APR-1999;	990S-0131149.	PR	21-JUL-1999;	990S-0144814.
PR	30-APR-1999;	990S-0132048.	PR	21-JUL-1999;	990S-0145086.
PR	30-APR-1999;	990S-0132407.	PR	21-JUL-1999;	990S-0145088.
PR	04-MAY-1999;	990S-0132484.	PR	22-JUL-1999;	990S-0145085.
PR	05-MAY-1999;	990S-0132485.	PR	22-JUL-1999;	990S-0145087.
PR	06-MAY-1999;	990S-0132486.	PR	22-JUL-1999;	990S-0145089.
PR	07-MAY-1999;	990S-0132487.	PR	22-JUL-1999;	990S-0145192.
PR	11-MAY-1999;	990S-0132863.	PR	23-JUL-1999;	990S-0145145.
PR	14-MAY-1999;	990S-0134218.	PR	23-JUL-1999;	990S-0145218.
PR	14-MAY-1999;	990S-0134219.	PR	23-JUL-1999;	990S-0145224.
PR	14-MAY-1999;	990S-0134221.	PR	26-JUL-1999;	990S-0145276.
PR	18-MAY-1999;	990S-0134370.	PR	27-JUL-1999;	990S-0145913.
PR	18-MAY-1999;	990S-0134768.	PR	27-JUL-1999;	990S-0145918.
PR	19-MAY-1999;	990S-0134941.	PR	27-JUL-1999;	990S-0145919.
PR	20-MAY-1999;	990S-0135124.	PR	28-JUL-1999;	990S-0145955.
PR	21-MAY-1999;	990S-0135353.	PR	02-AUG-1999;	990S-0146386.
PR	24-MAY-1999;	990S-0135629.	PR	02-AUG-1999;	990S-0146388.
PR	25-MAY-1999;	990S-0136021.	PR	02-AUG-1999;	990S-0146389.
PR	27-MAY-1999;	990S-0136392.	PR	03-AUG-1999;	990S-0147038.
PR	28-MAY-1999;	990S-0136782.	PR	04-AUG-1999;	990S-0147204.
PR	01-JUN-1999;	990S-0137222.	PR	04-AUG-1999;	990S-0147302.
PR	03-JUN-1999;	990S-0137528.	PR	05-AUG-1999;	990S-0147192.
PR	04-JUN-1999;	990S-0137502.	PR	05-AUG-1999;	990S-0147260.
PR	07-JUN-1999;	990S-0137724.	PR	06-AUG-1999;	990S-0147303.
PR	08-JUN-1999;	990S-0138094.	PR	06-AUG-1999;	990S-0147416.
PR	10-JUN-1999;	990S-0138340.	PR	09-AUG-1999;	990S-0147493.
PR	10-JUN-1999;	990S-0138340.	PR	09-AUG-1999;	990S-0147935.
PR	14-JUN-1999;	990S-0139119.	PR	10-AUG-1999;	990S-0148171.
PR	16-JUN-1999;	990S-0139452.	PR	11-AUG-1999;	990S-0148319.
PR	16-JUN-1999;	990S-0139453.	PR	12-AUG-1999;	990S-0148341.
PR	17-JUN-1999;	990S-0139492.	PR	13-AUG-1999;	990S-0148565.
PR	18-JUN-1999;	990S-0139454.	PR	13-AUG-1999;	990S-0148684.
PR	18-JUN-1999;	990S-0139455.	PR	16-AUG-1999;	990S-0149368.
PR	18-JUN-1999;	990S-0139456.	PR	17-AUG-1999;	990S-0149175.
PR	18-JUN-1999;	990S-0139457.	PR	18-AUG-1999;	990S-0149426.
PR	18-JUN-1999;	990S-0139458.	PR	20-AUG-1999;	990S-0149722.
PR	18-JUN-1999;	990S-0139459.	PR	20-AUG-1999;	990S-0149723.
PR	18-JUN-1999;	990S-0139460.	PR	20-AUG-1999;	990S-0149929.
PR	18-JUN-1999;	990S-0139461.	PR	20-AUG-1999;	990S-0149929.
PR	18-JUN-1999;	990S-0139462.	PR	23-AUG-1999;	990S-0149930.
PR	18-JUN-1999;	990S-0139463.	PR	25-AUG-1999;	990S-0150566.
PR	18-JUN-1999;	990S-0139750.	PR	26-AUG-1999;	990S-0150884.
PR	18-JUN-1999;	990S-0139763.	PR	27-AUG-1999;	990S-0151065.
PR	21-JUN-1999;	990S-0139817.	PR	27-AUG-1999;	990S-0151066.
PR	22-JUN-1999;	990S-0139899.	PR	27-AUG-1999;	990S-0151080.
PR	23-JUN-1999;	990S-0140353.	PR	30-AUG-1999;	990S-0151303.
PR	24-JUN-1999;	990S-0140354.	PR	31-AUG-1999;	990S-0151438.
PR	24-JUN-1999;	990S-0140695.	PR	16-SEP-1999;	990S-0154039.
PR	28-JUN-1999;	990S-			

Tue Oct 7 15:52:35 2003

```

PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      72.18; Score 31; DB 21; Length 423;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9
Db 117 FLAQGLVL 125

```

Search completed: October 7, 2003, 13:25:31  
 Job time : 15.9241 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:19:12 ; Search time 5.06962 Seconds  
(without alignments)  
170.726 Million cell updates/sec

Title: US-09-807-512-24  
Perfect score: 43  
Sequence: 1 FLMAQGAML 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues  
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_76:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	34	79.1	1037	B87254	translation initia
2	32	74.4	290	Ti8992	hypothetical prote
3	32	74.4	350	D65018	Ethanolamine opero
4	32	74.4	350	D81042	probable ARAC-type
5	32	74.4	350	G85886	probable ARAC-type
6	32	74.4	904	AF2304	polyA polymerase [
7	31	72.1	476	B84454	hypothetical prote
8	31	72.1	476	C84454	hypothetical prote
9	31	72.1	815	T00264	high carbon dioxid
10	31	72.1	3344	J01899	genome polyprotein
11	30	69.8	111	T05919	probable GDP disso
12	30	69.8	415	A82237	hypothetical prote
13	30	69.8	444	T01782	GDP dissociation i
14	30	69.8	501	F98050	lysosome (EC 3.2.1
15	30	69.8	544	S62527	probable amino-aci
16	30	69.8	546	T46718	probable farnesyl
17	30	69.8	595	B82231	transport ATP-bind
18	30	69.8	913	A82587	translation initia
19	30	69.8	913	B97369	translation initia
20	30	69.8	1339	T47841	hypothetical prote
21	30	69.8	2004	D88948	protein ZK1005.1 [
22	29	67.4	183	A83405	hypothetical prote
23	29	67.4	184	T39201	myosin regulatory
24	29	67.4	270	B83784	hypothetical prote
25	29	67.4	327	JT0964	cytochrome-c oxida
26	29	67.4	334	E95869	probable smc22-r p
27	29	67.4	346	T35786	transcription regu
28	29	67.4	367	A12606	hypothetical prote
29	29	67.4	367	H97388	periplasmic mannit

30	29	67.4	373	2	T49865	hypothetical prote
31	29	67.4	387	2	SW4522	hypothetical prote
32	29	67.4	392	1	RWHUPD	poliovirus recepto
33	29	67.4	392	2	B44194	poliovirus recepto
34	29	67.4	417	1	RWHUPA	poliovirus recepto
35	29	67.4	417	2	A44194	poliovirus recepto
36	29	67.4	447	2	S20711	neuraminidase - in
37	29	67.4	453	1	NMIV3	exo-alpha-sialidas
38	29	67.4	454	1	NMIV	exo-alpha-sialidas
39	29	67.4	468	1	NMIVAA	exo-alpha-sialidas
40	29	67.4	469	2	D84454	hypothetical prote
41	29	67.4	470	1	NMIVU7	exo-alpha-sialidas
42	29	67.4	470	1	NMIVAK	exo-alpha-sialidas
43	29	67.4	470	1	NMIVXL	exo-alpha-sialidas
44	29	67.4	470	2	S04801	exo-alpha-sialidas
45	29	67.4	677	2	F95232	immunity protein,

ALIGNMENTS

RESULT 1

B87254  
translation initiation factor IF-2 [imported] - Caulobacter crescentus  
C:Species: Caulobacter crescentus  
C>Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 10-May-2001  
C:Accession: B87254  
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, N.; J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of Caulobacter crescentus.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: B87254  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1037 <STO>  
C:Cross-references: GB:AE005673; NID:g13421134; PIDN:AAK2030.1; GSPDB:GN00148  
C:Genetics:  
A:Gene: CC0042  
C:Superfamily: translation initiation factor IF-2; translation elongation factor Tu h

Query Match 79.1%; Score 34; DB 2; Length 1037;  
Best Local Similarity 77.8%; Pred. No. 31;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9  
DB 479 FLMQGVML 487

RESULT 2

Ti8992  
hypothetical protein C06B8.6 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
C:Accession: T18992  
R:Steward, C.  
submitted to the EMBL Data Library, November 1996  
A:Reference number: Z19057  
A:Accession: T18992  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-230 <WIL>  
A:Cross-references: EMBL:Z81463; PIDN:CAB03849.1; GSPDB:GN00023; CESP:C06B8.6  
A:Experimental source: clone C06B8  
C:Genetics:  
A:Gene: CESP:C06B8.6  
A:Map position: 5  
A:introns: 63/1; 124/3; 194/2; 256/3  
C:Superfamily: Caenorhabditis hypothetical protein C49G7.2

Query Match 74.4%; Score 32; DB 2; Length 290;

Db	214	LMAMGAML 221										
RESULT 5												
G85886 probable ARAC-type regulatory protein yfeG [imported] - Escherichia coli (strain O157												
C:Species: Escherichia coli												
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001												
C:Accession: G85886												
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May												
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda												
Nature 409, 529-533, 2001												
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.												
A:Reference number: A85480; MUID:21074935; PMID:11206551												
A:Accession: G85886												
A:Status: preliminary												
A:Molecule type: DNA												
A:Residues: 1-350 <STO>												
A:Cross-references: GB:AE005174; NID:g12516818; PIDN:AAG57555.1; GSPDB:GN00145; UWGP:												
A:Experimental source: strain O157:H7, substrain EDL933												
C:Genetics:												
A:Gene: yfeG												
Query Match 74.4%; Score 32; DB 2; Length 350;												
Best Local Similarity 87.5%; Pred. No. 28;												
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;												
QY	2	LMQAQAML 9										
Db	214	LMAMGAML 221										
RESULT 6												
AF2304 polyA polymerase [imported] - Nostoc sp. (strain PCC 7120)												
C:Species: Nostoc sp. PCC 7120												
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120												
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002												
C:Accession: AF2304												
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu												
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata												
DNA Res. 8, 205-213, 2001												
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium												
A:Reference number: AB1807; MUID:21595285; PMID:11759840												
A:Accession: AF2304												
A:Status: preliminary												
A:Molecule type: DNA												
A:Residues: 1-904 <KUR>												
A:Cross-references: GB:BA000019; PIDN:BA875688.1; PID:g17133124; GSPDB:GN00179												
A:Experimental source: strain PCC 7120												
C:Genetics:												
A:Gene: all3989												
Query Match 74.4%; Score 32; DB 2; Length 904;												
Best Local Similarity 77.8%; Pred. No. 74;												
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;												
QY	1	FLMAQGAFL 9										
Db	168	WLMAQGAFL 176										
RESULT 7												
B84454 hypothetical protein At2g04050 [imported] - Arabidopsis thaliana												
C:Species: Arabidopsis thaliana (mouse-ear cress)												
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001												
C:Accession: B84454												
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.												
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon												
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter												
Nature 402, 761-768, 1999												
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.												



A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: B84454  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-476 <STO>  
A:Cross-references: GB:AE002093; NID:g4734009; PIDN:AAD28686.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: At2g04050  
A:Map position: 2

Query Match 72.1%; Score 31; DB 2; Length 476;  
Best Local Similarity 66.7%; Pred. No. 63;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9  
||:||||:  
Db 170 FLAQGLVL 178

## RESULT 8

C84454  
hypothetical protein At2g04070 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C:Accession: C84454  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: C84454  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-476 <STO>  
A:Cross-references: GB:AE002093; NID:g4734008; PIDN:AAD28685.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: At2g04070  
A:Map position: 2

Query Match 72.1%; Score 31; DB 2; Length 476;  
Best Local Similarity 66.7%; Pred. No. 63;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9  
||:||||:  
Db 170 FLAQGLVL 178

## RESULT 9

T00264  
high carbon dioxide response protein 2 - Chlorococcum littorale  
N:Alternate names: HCR2 protein  
C:Species: Chlorococcum littorale  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 21-Jul-2000  
C:Accession: T00264  
R:Sasaki, T.; Kurano, N.; Miyachi, S.  
Plant Cell Physiol. 39, 131-138, 1998  
A:Title: Cloning and characterization of high-CO2-specific cDNAs from a marine microalgae  
A:Reference number: Z14135; MUID:98220314; PMID:9559558  
A:Accession: T00264  
A>Status: translated  
A:Molecule type: mRNA  
A:Residues: 1-815 <SAS>  
A:Cross-references: EMBL:AB007993; NID:g2541871; PIDN:BAA22845.1; PID:g2541872

Query Match 72.1%; Score 31; DB 2; Length 815;  
Best Local Similarity 77.8%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9  
||:||||:  
Db 681 FCMAVGAML 689

## RESULT 10

JQ1899  
genome polyprotein - papaya ringspot virus  
N:Contains: 46K protein; 63K protein; 6K protein; coat protein; coat protein; cylindrical inclusion  
C:Species: papaya ringspot virus, PRSV  
C:Date: 30-Sep-1993 #sequence\_revision 20-Aug-1994 #text\_change 02-Feb-2001  
C:Accession: JQ1899; PQ0509; A56603; S24785; S35725  
R:Yeh, S.D.; Jan, F.J.; Chiang, C.H.; Doong, T.J.; Chen, M.C.; Chung, P.H.; Bau, H.J.  
J. Gen. Virol. 73, 2531-2541, 1992  
A:Title: Complete nucleotide sequence and genetic organization of papaya ringspot vi  
A:Reference number: JQ1899; MUID:93019006; PMID:1402799  
A:Accession: JQ1899  
A:Molecule type: genomic RNA  
A:Residues: 1-3344 <YEH1>  
A:Cross-references: GB:S46722; NID:g258106; PIDN:AAB23789.1; PID:g258107  
A:Accession: PQ0509  
A:Molecule type: protein  
A:Residues: 548-554; 1402-1410 <YEH2>  
R:Wang, C.H.; Yeh, S.D.  
Arch. Virol. 127, 345-354, 1992  
A:Title: Nucleotide sequence comparison of the 3'-terminal regions of severe, mild, a  
A:Reference number: A56603; MUID:93090098; PMID:1456896  
A:Accession: A56603  
A:Molecule type: genomic RNA  
A:Residues: 2561-3344 <WAN>  
A:Cross-references: EMBL:X67672; NID:g62378; PIDN:CAA47904.1; PID:g62379  
A:Experimental source: type P, strain HA  
C:Superfamily: tobacco etch virus genome polyprotein  
C:Keywords: ATP; coat protein; cylindrical inclusion protein; genome-linked protein;  
F:1-547/Product: 63K protein #status predicted <PRO>  
F:548-1004/Product: helper component-proteinase #status predicted <HCP>  
F:1005-1401/Product: 46K protein #status predicted <PRT>  
F:1402-2036/Product: cylindrical inclusion protein #status predicted <CIP>  
F:1486-1493/Region: nucleotide-binding motif A (P-loop)  
F:1571-1576/Region: nucleotide-binding motif B  
F:1575-1578/Region: DEXH motif  
F:2037-2093/Product: 6K protein #status predicted <PRI>  
F:2094-2520/Product: nuclear inclusion a protein #status predicted <NIA>  
F:2521-3037/Product: nuclear inclusion b protein #status predicted <NIB>  
F:3038-3344/Product: coat protein #status predicted <COP>  
F:2156/Binding site: phosphoryl-RNA (Tyr) (covalent) #status predicted

Query Match 72.1%; Score 31; DB 2; Length 3344;  
Best Local Similarity 77.8%; Pred. No. 4.8e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9  
||:||||:  
Db 2068 FTMAGGAML 2076

## RESULT 11

T05919  
probable GDP dissociation inhibitor protein - barley (fragment)  
C:Species: Hordeum vulgare (barley)  
C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 20-Jun-2000  
C:Accession: T05919  
R:Hess, W.R.; Golz, R.R.; Boerner, T.  
Plant Sci. 133, 191-201, 1998  
A:Title: Analysis of randomly selceted cDNAs reveals the expression of stress- and  
A:Reference number: Z15411  
A:Accession: T05919  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 1-111 <HES>  
A:Cross-references: EMBL:AJ000240; PIDN:CAA03964.1  
A:Experimental source: cv Haisa, leaf  
C:Superfamily: human GDP dissociation inhibitor XAP-4

Query Match 69.8%; Score 30; DB 2; Length 111;  
Best Local Similarity 55.6%; Pred. No. 23;

Tue Oct 7 15:52:37 2003

us-09-807-512-24.rpr

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9  
I:|I|I|I|:  
Db 78 FMWANGALV 86

RESULT 12

AE2237  
hypothetical protein alr3452 [imported] - Nostoc sp. (strain PCC 7120)  
C:Species: Nostoc sp. PCC 7120  
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
C:Accession: AE2237  
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AE2237  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-415 <KUR>  
A:Cross-references: GB:BA000019; PIDN:BA075151.1; PID:g17132585; GSPDB:GN00179  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: alr3452

Query Match 69.8%; Score 30; DB 2; Length 415;  
Best Local Similarity 75.0%; Pred. No. 92;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LMAQGAML 9  
I:|I|I|I|:  
Db 78 LLARGAML 85

RESULT 13

T01782  
GDP dissociation inhibitor - common tobacco  
C:Species: Nicotiana tabacum (common tobacco)  
C:Date: 19-Feb-1999 #sequence\_revision 19-Feb-1999 #text\_change 26-May-2000  
C:Accession: T01782  
R:Ezaki, B.; Koyanagi, M.; Gardner, R.C.; Matsumoto, H.  
submitted to the EMBL Data Library, July 1997  
A:Reference number: Z14424  
A:Accession: T01782  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-444 <EZA>  
A:Cross-references: EMBL:AF012823; NID:g2501849; PID:g2501850  
A:Experimental source: cultivar Samsun  
C:Genetics:  
A:Gene: GDI  
A:Function:  
C:Description: inhibits dissociation of GDP from GTP binding proteins  
C:Superfamily: human GDP dissociation inhibitor XAP-4

Query Match 69.8%; Score 30; DB 2; Length 444;  
Best Local Similarity 55.6%; Pred. No. 98;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9  
I:|I|I|I|:  
Db 79 FMWANGALV 87

RESULT 14

F98050  
lysozyme (EC 3.2.1.17) [imported] - Streptococcus pneumoniae (strain R6)  
C:Species: Streptococcus pneumoniae  
C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 02-Nov-2001  
C:Accession: F98050

R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.;  
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.  
Y, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.  
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
A:Reference number: A97872; MUID:21429245; PMID:11544234  
A:Accession: F98050  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-501 <KUR>  
A:Cross-references: GB:AE007317; PIDN:AAL00235.1; PID:g15459086; GSPDB:GN00174  
C:Genetics:  
A:Gene: lytC  
C:Keywords: glycosidase; hydrolase

Query Match 69.8%; Score 30; DB 2; Length 501;  
Best Local Similarity 66.7%; Pred. No. 11e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9  
I:|I|I|I|:  
Db 181 FLNGQGAMI 189

RESULT 15

S62527  
probable amino-acid permease - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 16-May-1996 #sequence\_revision 13-Mar-1997 #text\_change 10-Dec-1999  
C:Accession: T41435; T39133; S62527  
R:Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.  
submitted to the EMBL Data Library, October 1998  
A:Reference number: Z21993  
A:Accession: T41435  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-544 <SEE>  
A:Cross-references: EMBL:AL032824; PIDN:CAB37426.1; GSPDB:GN00068; SPDB:SPCC584.13  
A:Experimental source: strain 972h-; cosmid c584  
R:Lye, G.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.  
submitted to the EMBL Data Library, November 1995  
A:Reference number: Z21830  
A:Accession: T39133  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-544 <LYE>  
A:Cross-references: EMBL:Z66569; NID:g1052533; PIDN:CAA91521.1; PID:g1052542; SPDB:SP

Query Match 69.8%; Score 30; DB 2; Length 544;  
Best Local Similarity 55.6%; Pred. No. 1.2e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9  
I:|I|I|I|:  
Db 343 FMGGQGMV 351

Search completed: October 7, 2003, 13:32:38  
Job time : 7.06962 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:05:57 ; Search time 2.62025 seconds  
(without alignments)  
161.527 Million cell updates/sec

Title: US-09-807-512-24  
Perfect score: 43  
Sequence: 1 FLMAQGAWL 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	32	74.4	350	1 EUTR_ECOLI	P36547 escherichia
2	31	72.1	344	1 POLG_PRSVH	Q01901 p genome po
3	30	69.8	544	1 YC9D_SCHPO	Q09887 schizosacch
4	30	69.8	1068	1 HMDH_ASPTT	Q09782 aspergillus
5	29	67.4	187	1 MLR1_SCHPO	Q9uug5 schizosacch
6	29	67.4	197	1 ISP2_PSEPK	P59364 pseudomonas
7	29	67.4	298	1 HEMK_PASMU	Q9cn82 pasteurilla
8	29	67.4	327	1 COX2_SYNVU	P98054 synechococc
9	29	67.4	417	1 PVR_CERAE	P32506 cercopitheci
10	29	67.4	417	1 PVR_HUMAN	P15151 homo sapien
11	29	67.4	453	1 NRAM_IATWIL	P03470 influenza a
12	29	67.4	454	1 NRAM_IAPUE	P03468 influenza a
13	29	67.4	468	1 NRAM_TASH2	P03478 influenza a
14	29	67.4	470	1 NRAM_TACHI	P11485 influenza a
15	29	67.4	470	1 NRAM_TAKIE	P31348 influenza a
16	29	67.4	470	1 NRAM_IATLEN	P31349 influenza a
17	29	67.4	470	1 NRAM_IATUSS	P03469 influenza a
18	29	67.4	811	1 SYEB_OCEIH	Q8eph5 oceanobacil
19	29	67.4	3133	1 HMCT_BOMMO	P98092 bombyx mori
20	28	65.1	366	1 TORY_ECO57	P58361 escherichia
21	28	65.1	371	1 DFB3_TREPA	Q83048 treponema p
22	28	65.1	373	1 SNE_HAEIN	P43862 haemophilus
23	28	65.1	525	1 NACL_ANASP	Q8yq78 anabaena sp
24	28	65.1	558	1 TIMH_METJA	Q60297 methanococc
25	28	65.1	587	1 ASB2_HUMAN	Q96q27 homo sapien
26	28	65.1	1002	1 DOR_DROME	Q24314 drosophila
27	28	65.1	1067	1 IMB4_SCHPO	Q60100 schizosacch
28	28	65.1	1202	1 ALAA_ARATH	Q91i83 arabidopsis
29	27	62.8	94	1 FLIQ_TREPA	P74931 treponema p
30	27	62.8	112	1 VFA4_BPV2	P13301 bovine papil
31	27	62.8	120	1 V54_BPV1	P06926 bovine papil
32	27	62.8	141	1 V223_FOWPV	Q91512 fowlpox vir
33	27	62.8	155	1 PLAS_HORVU	P08248 hordeum vul

RESULT 1  
EUTR\_ECOLI  
ID EUTR\_ECOLI STANDARD; PRT; 350 AA.  
AC P36547: P77451;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Ethanolamine operon regulatory protein.  
GN EUTR OR B2437.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K12;  
RX MEDLINE=94131946; PubMed=8300522;  
RA Troup B., Jahn M., Hungerer C., Jahn D.;  
RT "Isolation of the hemf operon containing the gene for the Escherichia coli aerobic coproporphyrinogen III oxidase by in vivo complementation of a yeast HEM13 mutant.";  
RL J. Bacteriol. 176:673-680(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K12 / MG1655;  
RX MEDLINE=97426617; PubMed=9278503;  
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.;  
RT "The complete genome sequence of Escherichia coli K-12.";  
RL Science 277:1453-1474(1997).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K12;  
RX MEDLINE=97349980; PubMed=9205837;  
RA Yamamoto Y., Aiba H., Baba T., Hayashi K., Inada T., Isono K., Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N., Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H., Oshima T., Oyama S., Saito N., Sampei G., Satoh Y., Sivasubram S., Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C., Yamagata S., Horuchi T.;  
RT "Construction of a contiguous 874-kb sequence of the Escherichia coli K-12 genome corresponding to 50.0-68.8 min on the linkage map and analysis of its sequence features.";  
RL DNA Res. 4:91-113(1997).  
CC -!- FUNCTION: ACTIVATES THE TRANSCRIPTION OF THE EUT OPERON. ALSO POSITIVELY REGULATES ITS OWN TRANSCRIPTION. PROBABLY BINDS ETHANOLAMINE AND VITAMIN B12 AS EFFECTORS (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE ARAC/XVLS FAMILY OF TRANSCRIPTIONAL REGULATORS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

34 27 62.8 207 1 Y101\_CVMA5 P18453 murine coro  
35 27 62.8 222 1 UNG\_PASMU P57807 pasteurilla  
36 27 62.8 231 1 PSDA\_MOUSE Q92x22 mus musculus  
37 27 62.8 265 1 FDHD\_STAAM Q99fz2 staphylococ  
38 27 62.8 265 1 FDHD\_STAAM Q8nv95 staphylococ  
39 27 62.8 275 1 Y760\_METJA Q58170 methanococc  
40 27 62.8 283 1 KC2E\_ARATH Q80507 arabidopsis  
41 27 62.8 309 1 CLT2\_MOUSE Q920a1 mus musculus  
42 27 62.8 316 1 LXHA\_DICDI P52922 dictyostell  
43 27 62.8 332 1 AXHA\_ASPNG P79019 aspergillus  
44 27 62.8 338 1 SP54\_BACSU P40869 bacillus su  
45 27 62.8 351 1 RNFD\_YERPE Q8zed2 yersinia pe

ALIGNMENTS

```
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X75413; CAA53168.1; -.
DR EMBL; AE000331; AAC75490.1; -.
DR EMBL; D90873; BAA16320.1; -.
DR EMBL; D90874; BAA16326.1; -.
DR PIR; D65018; D65018.
DR EcoGene; EG12190; eutr.
DR InterPro; IPR000005; HTHArac.
DR Pfam; PF00165; HTH_Arac; 2.
DR SMART; SM00342; HTH_ARAC; 1.
DR PROSITE; PS00041; HTH_ARAC_FAMILY_1; 1.
DR PROSITE; PS01124; HTH_ARAC_FAMILY_2; 1.
KW Transcription regulation; DNA-binding; Activator; Complete proteome.
FT DNA_BIND 259 278 H-N-H MOTIF (BY SIMILARITY).
FT CONFLICT 276 276 A -> R (IN REF. 1).
FT CONFLICT 285 285 P -> R (IN REF. 1).
FT CONFLICT 285 285 P -> R (IN REF. 1).
FT CONFLICT 285 285 P -> R (IN REF. 1).
SQ SEQUENCE 350 AA; 40160 MW; 1281A43DC31DE845 CRC64;

Query Match 74.4%; Score 32; DB 1; Length 350;
Best Local Similarity 87.5%; Pred. No. 9.5;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LMAQGAML 9
Db 214 LMAGAML 221

RESULT 2
POLG_PRSVH STANDARD; PRT; 3344 AA.
AC Q01901;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: N-terminal protein (P1); Helper
DE component proteinase (EC 3.4.22.45) (HC-Pro); Protein P3; 6 kDa
DE protein 1 (6k1); Cytoplasmic inclusion protein (CI); 6 kDa protein 2
DE (6k2); Genome-linked protein (VPG); Nuclear inclusion protein A (NI-A)
DE (NTA) (EC 3.4.22.44) (49 kDa proteinase) (49 kDa-Pro); Nuclear
DE inclusion protein B (NI-B) (NIB) (RNA-directed RNA polymerase)
DE (EC 2.7.7.48); Coat protein (CP)].
OS Papaya ringspot virus (strain P / mutant HA).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyvirus.
OC NCBI_TaxID=31731;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang C.H., Bau H.J., Yeh S.D.;
RT "Comparison of the nuclear inclusion b protein and coat protein genes
RT of five papaya ringspot virus strains distinct in geographic origin
RT and pathogenicity.";
RL Phytopathology 84:1205-1210(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93019006; PubMed=1402799;
RA Yeh S.D., Jan F.J., Chiang C.H., Doong T.J., Chen M.C.,
RA Chung P.H., Bau H.J.;
RT "complete nucleotide sequence and genetic organization of papaya
RT ringspot virus RNA.";
RL J. Gen. Virol. 73:2531-2541(1992).
RN [3]
RP SEQUENCE OF 2561-3344 FROM N.A.
RX MEDLINE=93090098; PubMed=1456896;
RA Wang C.H., Yeh S.D.;
RT "Nucleotide sequence comparison of the 3'-terminal regions of severe,
RT mild, and non-papaya infecting strains of papaya ringspot virus.";
RL Arch. Virol. 127:345-354(1992).
CC -1- FUNCTION: HELPER COMPONENT-PROTEINASE IS REQUIRED FOR APHID
CC TRANSMISSION AND ALSO HAS PROTEOLYTIC ACTIVITY.
CC -1- FUNCTION: CYTOPLASMIC INCLUSION PROTEIN HAS HELICASE ACTIVITY. IT
```

```

FT BINDING      2156      2156      COVALENT LINKAGE OF VIRAL RNA (BY
FT NP_BIND      1486      1493      ATP (POTENTIAL);
FT SEQUENCE     3344 AA; 381040 MW; E90CD7523AC5243D CRC64;
SQ
Query Match      72.1%; Score 31; DB 1; Length 3344;
Best Local Similarity 77.8%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 FLMAOGAML 9
Db      2068 FTMAGGAML 2076

RESULT 3
YC9D_SCHPO
ID YC9D_SCHPO      STANDARD;      PRT;      544 AA.
AC Q09887;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative amino-acid permease C584.i3.
GN SPCC584.i3.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OC NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer C., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie M.R., Paulsen I., Potashkin J.,
RA Shpakovsky G.V., Ussery D., Barrell B.G., Nurse P.;
RT The genome sequence of Schizosaccharomyces pombe.;
RL Nature 415:871-880(2002).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: Belongs to the amino acid permease family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL032824; CAB37426.1;
CC PIR; T41435; S62527.
CC GenBank; SPombe; SPCC584.i3;
CC InterPro; IPR002293; Aa/permease1.

```

```

DR InterPro; IPR004840; AAC_permease.
DR InterPro; IPR004841; Permease.
DR Pfam; PF00324; aa_permeases; 1.
DR PROSITE; PS00218; AMINO_ACID_PERMEASE_1; FALSE_NEG.
KW Hypothetical protein; Transport; Amino-acid transport; Transmembrane.
FT TRANSMEM 41 61 POTENTIAL.
FT TRANSMEM 71 91 POTENTIAL.
FT TRANSMEM 106 126 POTENTIAL.
FT TRANSMEM 166 186 POTENTIAL.
FT TRANSMEM 196 216 POTENTIAL.
FT TRANSMEM 240 260 POTENTIAL.
FT TRANSMEM 280 300 POTENTIAL.
FT TRANSMEM 332 352 POTENTIAL.
FT TRANSMEM 391 411 POTENTIAL.
FT TRANSMEM 413 433 POTENTIAL.
FT TRANSMEM 451 471 POTENTIAL.
FT TRANSMEM 484 504 POTENTIAL.
SQ SEQUENCE 544 AA; 59527 MW; 39E4C220F63BA3CB CRC64;

Query Match      69.8%; Score 30; DB 1; Length 544;
Best Local Similarity 55.6%; Pred. No. 42;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 FLMAOGAML 9
Db      343 FMGOGCMV 351

RESULT 4
HMDH_ASPTD
ID HMDH_ASPTD      STANDARD;      PRT;      1068 AA.
AC Q9Y7D2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 3-hydroxy-3-methylglutaryl-coenzyme A reductase (EC 1.1.1.34) (HMG-CoA
DE reductase).
OS Aspergillus terreus.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OC NCBI_TaxID=33178;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 20542;
RX MEDLINE=99269277; PubMed=10334994;
RA Kennedy J., Auclair K., Kendrew S.G., Park C., Vederas J.C.,
RA Hutchinson C.R.;
RT "Modulation of polyketide synthase activity by accessory proteins
RT during lovastatin biosynthesis.";
RL Science 284:1368-1372(1999).
CC -!- FUNCTION: INVOLVED IN THE CONTROL OF CHOLESTEROL BIOSYNTHESIS. IT
CC IS THE RATE-LIMITING ENZYME OF THE STEROL BIOSYNTHESIS.
CC -!- CATALYTIC ACTIVITY: (R)-mevalonate + CoA + 2 NADP(+) -> (S)-3-
CC hydroxy-3-methylglutaryl-CoA + 2 NADPH.
CC -!- PATHWAY: Cholesterol biosynthesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC reticulum.
CC -!- SIMILARITY: BELONGS TO THE HMG-COA REDUCTASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF141925; AAD34556.1;
CC InterPro; IPR002202; HMG-CoA_red.
CC InterPro; IPR004554; HMG-CoA_R_NADP.
CC Pfam; IPR000731; SSD_5TM.
CC Pfam; PF00368; HMG-CoA_red; 1.
CC PRINTS; PR00071; HMGCOARDTASE.

```

DR	TIGRFAMS; TIGR00533; HMG-CoA_R_WADP; 1.	RA	Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
DR	PROSITE; PS00066; HMG-CoA_REDUCTASE_1; 1.	RA	Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
DR	PROSITE; PS00318; HMG-CoA_REDUCTASE_2; 1.	RA	Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
DR	PROSITE; PS01192; HMG-CoA_REDUCTASE_3; 1.	RA	Cerrutti L., Lowe I., McCombie W.R., Paulsen I., Potashkin J.,
DR	PROSITE; PS00065; HMG-CoA_REDUCTASE_4; 1.	RA	Shpakovski G.V., ussery D., Barrell B.G., Nurse P.,
DR	PROSITE; PS0156; SSD; 1.	RT	"The genome sequence of Schizosaccharomyces pombe";
KW	Oxidoreductase; Glycoprotein; Endoplasmic reticulum; Transmembrane;	RL	Nature 415:871-880(2002).
KW	Cholesterol biosynthesis; NADP.	RL	[2]
FT	DOMAIN 1 537	RN	IDENTIFICATION AND CHARACTERIZATION.
FT	DOMAIN 538 588	RX	MEDLINE=205223843; PubMed=11069761;
FT	DOMAIN 589 1059	RA	Le Goff X., Motejl F., Salimova E., Mabuchi I., Simanis V.;
FT	TRANSMEM 195 215	RA	"The S. pombe rlc1 gene encodes a putative myosin regulatory light
FT	TRANSMEM 225 245	RT	chain that binds the type II myosins myo3p and myo2p.";
FT	POTENTIAL 251 271	RL	J. Cell Sci. 113:4157-4163(2000).
FT	POTENTIAL 251 271	CC	-!- SUBUNIT: BINDS TO MYOSIN II CHAINS MYO2 AND MYO3.
FT	TRANSMEM 352 372	CC	-!- SUBCELLULAR LOCATION: Cytoplasmic.
FT	POTENTIAL 414 434	CC	-!- MISCELLANEOUS: THIS CHAIN BINDS CALCIUM (Potential).
FT	POTENTIAL 517 537	CC	-!- SIMILARITY: Contains 1 EF-hand calcium-binding domain.
FT	TRANSMEM 703 703	CC	-----
FT	ACT_SITE 913 913	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
FT	ACT_SITE 1009 1009	CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -
FT	CARBOHYD 410 410	CC	the European Bioinformatics Institute. There are no restrictions on its
FT	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	use by non-profit institutions as long as its content is in no way
FT	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	modified and this statement is not removed. Usage by and for commercial
FT	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
FT	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	or send an email to license@isb-sib.ch).
FT	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	-----
FT	CARBOHYD 1013 1013	DR	EMBL; AL110469; CAB54151.1; -.
FT	CARBOHYD 1013 1013	DR	PIR; T39201; T39201.
FT	SEQUENCE 1068 AA; 6082140D839CA271 CRC64;	DR	HSSP; P02593; ICDM.
		DR	GeneDB_SPombe; SPAC926.03; -.
		DR	InterPro; IPR002048; EF-hand.
		DR	Pfam; PF00036; efhnd; 2.
		DR	ProDom; PD000012; EF-hand; 1.
		DR	SMART; SM00054; EFh; 2.
		DR	PROSITE; PS00018; EF-HAND; 1.
		KW	Myosin; Calcium-binding.
		FT	CA_BIND 58 69
		FT	SEQUENCE 184 AA; 19997 MM; 0E7CD52BF74ADD1E CRC64;
		QY	Query Match 67.4%; Score 29; DB 1; Length 184;
		DB	Best Local Similarity 66.7%; Pred. No. 25;
			Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
			QY 1 FLMAQGAML 9
		DB	102 FLTAMGSM 110
			-----
			RESULT 6
			ISP2_PSEPK
			ID. ISP2_PSEPK STANDARD; PRT; 197 AA.
			AC. P59364;
			DT 28-FEB-2003 (Rel. 41, Created)
			DT 28-FEB-2003 (Rel. 41, Last sequence update)
			DT 15-SEP-2003 (Rel. 42, Last annotation update)
			DE Probable intracellular septation protein.
			GN ISP2 OR P4501.
			OS Pseudomonas putida (strain KT2440).
			OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
			OC Pseudomonadaceae; Pseudomonas.
			OX NCBI_TaxID=160488;
			RN [1]
			RP MEDLINE=22423060; PubMed=12534463;
			RX Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
			RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
			RA Brinkup L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
			RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
			RA Chris Lee P., Holtzapfel E., Scanlan D., Tran K., Moazzes A.,
			RA Utterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,
			RA Lauber J., Stjepandic D., Honeisel J., Straetz M., Heim S.,
			RA Kiewitz C., Eisen J., Timmis K.N., Duesterhoeft A., Tuemmler B.,

RA Fraser C.M.;  
 RT "complete genome sequence and comparative analysis of the  
 RL metabolically versatile *Pseudomonas putida* KT2440.";  
 CC Environ. Microbiol. 4:799-808(2002).  
 CC -!- FUNCTION: Involved in cell division; probably involved in  
 CC intracellular septation (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
 CC -!- SIMILARITY: BELONGS TO THE ISPZ FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AE016790; AAN70075.1; -;  
 DR TIGR: PF4501; -; 1;  
 DR HAMAP: MF\_00189; -; 1;  
 DR Pfam: PF04279; ISPA; 1.  
 DR ProDom: PD016710; SeptationA; 1.  
 DR TIGRFAMS: TIGR00997; ispZ; 1.  
 KW Cell division; Septation; Transmembrane; Complete proteome.  
 FT TRANSMEM 33 55 POTENTIAL.  
 FT TRANSMEM 64 83 POTENTIAL.  
 FT TRANSMEM 93 115 POTENTIAL.  
 FT TRANSMEM 135 154 POTENTIAL.  
 FT TRANSMEM 164 186 POTENTIAL.  
 SQ SEQUENCE 197 AA; 22400 MW; F3107A5ECCD9A529 CRC64;  
 Query Match 67.4%; Score 29; DB 1; Length 197;  
 Best Local Similarity 66.7%; Pred. No. 26;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 FLMAOGAML 9  
 DB 174 FLVAGVYL 182  
 RESULT 7  
 ID HEMK\_PASMU STANDARD; PRT; 298 AA.  
 AC Q9CN82;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE HemK protein homolog (EC 2.1.1.-) (M.PnuHemRP).  
 GN HEMK OR PM0556.  
 OS Pasteurella multocida.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
 OC Pasteurellaceae; Pasteurella.  
 OX NCBI\_TaxID=747;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Pm70;  
 RX MEDLINE=21145866; PubMed=11248100;  
 RA May B.J., Zhang Q., Li L., Paustian M.L., Whittam T.S., Kapur V.;  
 RT "Complete genomic sequence of *Pasteurella multocida* Pm70.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).  
 CC -!- SIMILARITY: BELONGS TO THE HEMK FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AE006091; AAK02640.1; -;  
 DR REBASE: 5515; M.PnuHemKP.  
 DR InterPro: IPR004556; HemK.  
 DR

DR InterPro: IPR002052; N6\_Mtase.  
 DR InterPro: IPR000051; SAM\_bind.  
 DR TIGRFAMS: TIGR00536; hemK\_fam; 1.  
 DR PROSITE: PS00092; N6\_MTASE; UNKNOWN1.  
 KW Transferase; Methyltransferase; Complete proteome.  
 SQ SEQUENCE 298 AA; 33428 MW; 07624CC6EA2CC2D8 CRC64;  
 Query Match 67.4%; Score 29; DB 1; Length 298;  
 Best Local Similarity 66.7%; Pred. No. 39;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 FLMAOGAML 9  
 DB 249 FLKPGCAL 257  
 RESULT 8  
 ID COX2\_SYNNU STANDARD; PRT; 327 AA.  
 AC P98054;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Cytochrome c oxidase polypeptide II precursor (EC 1.9.3.1)  
 DE (Cytochrome AA3 subunit 2) (Oxidase AA(3) subunit 2).  
 GN CTAC.  
 OS *Synechococcus vulcanus* (Thermosynechococcus vulcanus).  
 OC Bacteria; Cyanobacteria; Chroococcales; Thermosynechococcus.  
 OX NCBI\_TaxID=32053;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92068230; PubMed=1659815;  
 RA Tano H., Sone N.;  
 RT "The cytochrome C oxidase genes in blue-green algae and  
 RT characteristics of the deduced protein sequence for subunit II of the  
 RL thermophilic cyanobacterium *Synechococcus vulcanus*.";  
 RL Biochem. Biophys. Res. Commun. 181:437-442(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94002240; PubMed=8399373;  
 RA Sone N., Tano H., Ishizuka M.;  
 RT "The genes in the thermophilic cyanobacterium *Synechococcus vulcanus*  
 RL encoding cytochrome-c oxidase.";  
 RL Biochim. Biophys. Acta 1183:130-138(1993).  
 CC -!- FUNCTION: SUBUNIT I AND II FORM THE FUNCTIONAL CORE OF THE ENZYME  
 CC COMPLEX. ELECTRONS ORIGINATING IN CYTOCHROME C ARE TRANSFERRED VIA  
 CC HEME A AND CU(A) TO THE BINUCLEAR CENTER FORMED BY HEME A3 AND  
 CC CU(B).  
 CC -!- CATALYTIC ACTIVITY: 4 ferrocyclochrome c + O(2) = 4 ferricytochrome  
 CC c + 2 H(2)O.  
 CC -!- COFACTOR: COPPER A.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: D16254; BAA41040.1; -;  
 DR InterPro: IPR001505; Copper\_CuA.  
 DR InterPro: IPR002429; Cyt\_c\_ox\_2.  
 DR Pfam: PF00116; COX2; 1.  
 DR Pfam: PF02790; COX2\_TM; 1.  
 DR ProDom: PD000131; Copper\_CuA; 1.  
 DR PROSITE: PS00078; COX2; 1.  
 KW Oxidoreductase; Respiratory chain; Electron transport; Transmembrane;  
 KW Copper; Signal.  
 FT SIGNAL 1 23 POTENTIAL.  
 FT CHAIN 24 327 CYTOCHROME C OXIDASE POLYPEPTIDE II.  
 FT

FT	TRANSMEM	56	78	POTENTIAL.
FT	TRANSMEM	96	114	POTENTIAL.
FT	METAL	221	221	COPPER A (PROBABLE).
FT	METAL	255	255	COPPER A (PROBABLE).
FT	METAL	259	259	COPPER A (PROBABLE).
FT	METAL	263	263	COPPER A (PROBABLE).
SQ	SEQUENCE	327 AA;	35355 MW;	2C28F9C68D3B7A13 CRC64;
Query Match 67.4%; Score 29; DB 1; Length 327;				
Best Local Similarity 55.6%; Pred. No. 43;				
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;				
Qy	1 FLMAQGAML 9			
Db	57 FLVVGAI 65			
RESULT 9				
ID	PVR_CERAE	STANDARD;	PRT;	417 AA.
AC	P32506;			
DT	01-OCT-1993 (Rel. 27, Created)			
DT	01-OCT-1993 (Rel. 27, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Poliovirus receptor precursor.			
GN	PVR OR PVS.			
OS	Cercopithecus aethiops (Green monkey) (Grivet).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Primates; Catarrhini; Cercopithecidae;			
OC	Cercopithecinae; Cercopithecus.			
OX	NCBI_TaxID=9534;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORMS ALPHA AND DELTA).			
RC	TISSUE=Kidney;			
RX	MEDLINE=93059651; PubMed=1331508;			
RA	Koike S., Ise I., Sato Y., Yonekawa H., Gotoh O., Nomoto A.;			
RT	"A second gene for the African green monkey poliovirus receptor that has no putative N-glycosylation site in the functional N-terminal immunoglobulin-like domain."			
RT	J. Virol. 66:7059-7066(1992).			
CC	- FUNCTION: NOT KNOWN. USED BY POLIOVIRUS TO BIND AND ENTER THE CELL.			
CC	- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (ALPHA FORM);			
CC	SECRETED (BETA AND GAMMA FORMS).			
CC	- ALTERNATIVE PRODUCTS:			
CC	Event-Alternative splicing; Named isoforms=4;			
CC	Name=Alpha;			
CC	ISOID=P32506-1; Sequence=Displayed;			
CC	Name=Beta;			
CC	ISOID=P32506-3; Sequence=Not described;			
CC	Name=Gamma;			
CC	ISOID=P32506-4; Sequence=Not described;			
CC	Name=Delta;			
CC	ISOID=P32506-2; Sequence=VSP_002622, VSP_002623;			
CC	- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.			
CC	- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.			
CC	- SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; D12611; BAA02136.1; -			
DR	EMBL; D12612; BAA02137.1; -			
DR	PIR; A44194; A44194.			
DR	PIR; B44194; B44194.			
DR	InterPro; IPR007110; Ig-like.			
DR	InterPro; IPR003006; Ig_MHC.			
DR	InterPro; IPR003596; Ig_V.			

DR	Pfam: PF00047; Ig; 3.			
DR	SMART; SM00406; IGV; 1.			
DR	PROSITE; PS00835; IG-LIKE; 3.			
KW	Immunoglobulin domain; Receptor; Transmembrane; Glycoprotein; Signal;			
KW	Repeat; Alternative splicing.			
FT	SIGNAL	1	20	POTENTIAL.
FT	CHAIN	21	417	POLIOVIRUS RECEPTOR.
FT	DOMAIN	21	343	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	344	367	POTENTIAL.
FT	DOMAIN	368	417	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	24	139	IG-LIKE V-TYPE.
FT	DOMAIN	145	237	IG-LIKE C2-TYPE 1.
FT	DOMAIN	244	328	IG-LIKE C2-TYPE 2.
FT	DISULFID	49	123	BY SIMILARITY.
FT	DISULFID	166	221	BY SIMILARITY.
FT	DISULFID	266	312	BY SIMILARITY.
FT	CARBOHYD	85	85	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	120	120	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	188	188	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	237	237	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	278	278	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	307	307	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	313	313	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARSPLIC	386	392	EHASASA -> HHOSCHN (in isoform Delta).
FT	VARSPLIC	393	417	/FTIG-VSP_002622.
FT	VARSPLIC	393	417	Missing (in isoform Delta).
FT	SEQUENCE	417 AA;	45464 MW;	DA4AD0FE4D2F6E1F CRC64;
SQ	SEQUENCE	417 AA;	45464 MW;	DA4AD0FE4D2F6E1F CRC64;
Query Match 67.4%; Score 29; DB 1; Length 417;				
Best Local Similarity 66.7%; Pred. No. 55;				
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;				
Qy	1 FLMAQGAML 9			
Db	289 FAVAQAQL 297			
RESULT 10				
ID	PVR_HUMAN	STANDARD;	PRT;	417 AA.
AC	P15151; P15152; Q15267; Q15268;			
DT	01-FEB-1990 (Rel. 14, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Poliovirus receptor precursor (CD155 antigen).			
GN	PVR OR PVS.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89168426; PubMed=2538245;			
RA	Mendelsohn C.L., Wimmer E., Racaniello V.R.;			
RT	"Cellular receptor for poliovirus: molecular cloning, nucleotide sequence, and expression of a new member of the immunoglobulin superfamily."			
RT	Cell 56:855-865(1989).			
RL	[2]			
RN	REVISIONS.			
RA	Racaniello V.R.;			
RL	Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.			
RP	[3]			
RN	SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.			
RX	MEDLINE=91006015; PubMed=2170108;			
RA	Koike S., Horie H., Ise I., Okitsu A., Yoshida M., Iizuka N.,			
RT	Takeuchi K., Takegami T., Nomoto A.;			
RT	"The poliovirus receptor protein is produced both as membrane-bound and secreted forms";			
RL	EMBO J. 9:3217-3224 (1990).			
RN	[4]			
RP	SEQUENCE FROM N.A.			



RA Kodyianni V., Severin J., Ge Y., Grable L., Kvistad E., Gordon L.,  
RA Shannon M., Brower A., Olsen A.S., Smith L.M.;  
RT "Sequence analysis of a lmb region in 19q13.2 containing a zinc finger  
RT gene cluster.";  
RT Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RN DOMAINS.  
RX MEDLINE=91239515; PubMed=1851992;  
RA Zibert A., Wimmer E.;  
RA "N glycosylation of the virus binding domain is not essential for  
RT function of the human poliovirus receptor.";  
RL J. Virol. 66:7368-7373(1992).  
RN [6]  
RN MUTAGENESIS OF CARBOHYDRATE-LINKAGE SITES.  
RX MEDLINE=93059689; PubMed=1331527;  
RA Zibert A., Wimmer E.;  
RA "N glycosylation of the virus binding domain is not essential for  
RT function of the human poliovirus receptor.";  
RL J. Virol. 66:7368-7373(1992).  
CC -!- FUNCTION: NOT KNOWN. USED BY POLIOVIRUS TO BIND AND ENTER THE  
CC CELL.  
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (ALPHA AND DELTA  
CC FORMS); SECRETED (BETA AND GAMMA FORMS).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event-Alternative splicing; Named isoforms=4;  
CC Name=Alpha; Sequence=Displayed;  
CC Name=Beta;  
CC IsoId=PI5151-1; Sequence=VSP\_002617;  
CC Name=Gamma;  
CC IsoId=PI5151-2; Sequence=VSP\_002618; VSP\_002619;  
CC IsoId=PI5151-3; Sequence=VSP\_002618, VSP\_002619;  
CC Name=Delta;  
CC IsoId=PI5151-4; Sequence=VSP\_002620, VSP\_002621;  
CC -!- MISCELLANEOUS: THE V-TYPE DOMAIN IS NECESSARY AND SUFFICIENT FOR  
CC VIRUS BINDING AND UPTAKE.  
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.  
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
CC -!- SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.  
CC -!- DATABASE: NAME=PROV; NOTE=Cd guide cd155 entry;  
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd155.htm".  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M24407; AAA36461.1; -  
CC EMBL; M24406; AAA36462.1; -  
CC EMBL; X64116; CAA45478.1; -  
CC EMBL; X64117; CAA45478.1; JOINED.  
CC EMBL; X64118; CAA45478.1; JOINED.  
CC EMBL; X64119; CAA45478.1; JOINED.  
CC EMBL; X64120; CAA45478.1; JOINED.  
CC EMBL; X64121; CAA45478.1; JOINED.  
CC EMBL; X64122; CAA45478.1; JOINED.  
CC EMBL; X64123; CAA45478.1; JOINED.  
CC EMBL; X64116; CAA45479.1; -  
CC EMBL; X64117; CAA45479.1; JOINED.  
CC EMBL; X64118; CAA45479.1; JOINED.  
CC EMBL; X64119; CAA45479.1; JOINED.  
CC EMBL; X64120; CAA45479.1; JOINED.  
CC EMBL; X64121; CAA45479.1; JOINED.  
CC EMBL; X64122; CAA45479.1; JOINED.  
CC EMBL; X64123; CAA45479.1; JOINED.  
CC EMBL; X64116; CAA45480.1; -  
CC EMBL; X64117; CAA45480.1; JOINED.  
CC EMBL; X64118; CAA45480.1; JOINED.  
CC EMBL; X64119; CAA45480.1; JOINED.  
CC EMBL; X64120; CAA45480.1; JOINED.  
CC EMBL; X64121; CAA45480.1; JOINED.  
CC EMBL; X64122; CAA45480.1; JOINED.  
CC EMBL; X64123; CAA45480.1; JOINED.

DR EMBL; AC068948; AAF69803.1; -  
DR PIR; S12048; RWHUPA.  
DR Genew; HGNC:9705; PVR.  
DR MIM; 173850; -  
DR GO; GO:0005737; C:cytoplasm; TAS.  
DR GO; GO:0005615; C:extracellular space; TAS.  
DR GO; GO:0016021; C:integral to membrane; TAS.  
DR GO; GO:0004872; F:receptor activity; TAS.  
DR GO; GO:0007125; P:invasive growth; TAS.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003596; Ig\_v.  
DR Pfam; PF00047; Ig; 3.  
DR SMART; SM00406; Igv; 1.  
DR PROSITE; PS50835; IG\_LIKE; 3.  
KW Immunoglobulin domain; Receptor; Transmembrane; Glycoprotein; Signal;  
KW Repeat; Antigen; Alternative splicing; Polymorphism.  
FT SIGNAL 1 20  
FT CHAIN 21 417  
FT DOMAIN 21 343  
FT TRANSMEM 344 367  
FT DOMAIN 368 417  
FT DOMAIN 24 139  
FT DOMAIN 145 237  
FT DOMAIN 244 328  
FT DISULFID 49 123  
FT DISULFID 166 221  
FT DISULFID 266 312  
FT CARBOHYD 105 105  
FT CARBOHYD 120 120  
FT CARBOHYD 188 188  
FT CARBOHYD 218 218  
FT CARBOHYD 237 237  
FT CARBOHYD 278 278  
FT CARBOHYD 307 307  
FT CARBOHYD 313 313  
FT VARSPPLIC 340 384  
FT VARSPPLIC 331 331  
FT VARSPPLIC 332 384  
FT VARSPPLIC 395 392  
FT VARSPPLIC 393 417  
FT VARIANT 67 67  
FT VARIANT 340 340  
SQ SEQUENCE 417 AA; 45302 MW; D15C012CE853169B CRC64;  
Query Match 67.4%; Score 29; DB 1; Length 417;  
Best Local Similarity 66.7%; Pred. No. 55;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 1 FLMAQAGML 9  
Db 289 FAVAQAGQL 297  
RESULT 11  
ID NRAM IAWIL STANDARD; PRT; 453 AA.  
AC P03470;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Neuraminidase (EC 3.2.1.18).  
GN NA.  
OS Influenza A virus (strain A/Wilson-Smith/33).  
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;  
OC Influenza A viruses; Influenzavirus A.

```

OX NCBI_TaxID=11487;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82192605; PubMed=7077751;
RA Hiti A.L., Nayak D.P.;
RT "Complete nucleotide sequence of the neuraminidase gene of human
RL Influenza virus A/WSN/33.";
CC J. Virol. 41:730-734(1982).
CC
CC -1- FUNCTION: Removes the terminal sialic acid from carbohydrate side
CC chains of the host cell surface proteins and from the viral
CC envelope. Such a reaction prevents self-aggregation and facilitate
CC the mobility of the virus to and from the site of infection.
CC
CC -1- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2->3)-, alpha-(2->6)-,
CC alpha-(2->8)-glycosidic linkages of terminal sialic residues in
CC oligosaccharides, glycoproteins, glycolipids, colominic acid and
CC synthetic substrates.
CC
CC -1- SUBUNIT: Homotrimer.
CC
CC -1- SUBCELLULAR LOCATION: VIRAL MEMBRANE. FORMS A MUSHROOM-SHAPED
CC SPIKE ON THE SURFACE OF THE VIRION.
CC
CC -1- SIMILARITY: BELONGS TO FAMILY 34 OF GLYCOSYL HYDROLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: J02177; AAA43397.1; ALT_SEQ.
CC HSSP: P03472; 20WC.
CC InterPro: IPR001860; Glyco_hydro_34.
CC Pfam: PF00064; neur; 1.
CC ProDom: PD000431; Glyco_hydro_34; 1.
CC Hydrolase; Glycosidase; Glycoprotein; Transmembrane.
CC TRANSMEM 7 35
CC DOMAIN 36 74 HYPERVARIABLE STALK REGION.
CC ACT_SITE 259 259 PROBABLE.
CC CARBOHYD 261 261 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 44 44 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 72 72 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 219 219 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 382 382 N-LINKED (GLCNAC. .) (POTENTIAL).
CC SEQUENCE 453 AA; 49623 MW; 7DC56A4416A47BE8 CRC64;
CC
CC Query Match 67.4%; Score 29; DB 1; Length 453;
CC Best Local Similarity 55.6%; Pred. No. 60;
CC Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
CC
CC QY 1 FLMAQGAML 9
CC Db 116 FFLTQGALL 124
CC
CC RESULT 12
CC NRAM_IAPUE STANDARD; PRT; 454 AA.
CC AC P03468;
CC DT 21-JUL-1986 (Rel. 01, Created)
CC DT 21-JUL-1986 (Rel. 01, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Neuraminidase (EC 3.2.1.18).
CC GN NA.
CC OS Influenza A virus (strain A/Puerto Rico/8/34).
CC OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
CC OC Influenza A viruses; Influenzavirus A.
CC OX NCBI_TaxID=11455;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RX MEDLINE=81148841; PubMed=7010182;
CC RA Fields S., Winter G., Brownlee G.G.;
CC RT "Structure of the neuraminidase gene in human influenza virus

```

```

RT A/PR/8/34.";
RL Nature 290:213-217(1981).
CC
CC -1- FUNCTION: Removes the terminal sialic acid from carbohydrate side
CC chains of the host cell surface proteins and from the viral
CC envelope. Such a reaction prevents self-aggregation and facilitate
CC the mobility of the virus to and from the site of infection.
CC
CC -1- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2->3)-, alpha-(2->6)-,
CC alpha-(2->8)-glycosidic linkages of terminal sialic residues in
CC oligosaccharides, glycoproteins, glycolipids, colominic acid and
CC synthetic substrates.
CC
CC -1- SUBUNIT: Homotrimer.
CC
CC -1- SUBCELLULAR LOCATION: VIRAL MEMBRANE. FORMS A MUSHROOM-SHAPED
CC SPIKE ON THE SURFACE OF THE VIRION.
CC
CC -1- SIMILARITY: BELONGS TO FAMILY 34 OF GLYCOSYL HYDROLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: J02146; AAA43412.1; -.
CC HSSP: P03472; 2QWC.
CC InterPro: IPR001860; Glyco_hydro_34.
CC Pfam: PF00064; neur; 1.
CC ProDom: PD000431; Glyco_hydro_34; 1.
CC Hydrolase; Glycosidase; Glycoprotein; Transmembrane.
CC TRANSMEM 7 35
CC DOMAIN 36 75 HYPERVARIABLE STALK REGION.
CC ACT_SITE 260 260 PROBABLE.
CC CARBOHYD 44 44 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 58 58 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 73 73 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 131 131 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 220 220 N-LINKED (GLCNAC. .) (POTENTIAL).
CC SEQUENCE 454 AA; 50143 MW; A0DC4C08A2B53705 CRC64;
CC
CC Query Match 67.4%; Score 29; DB 1; Length 454;
CC Best Local Similarity 55.6%; Pred. No. 60;
CC Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
CC
CC QY 1 FLMAQGAML 9
CC Db 117 FFLTQGALL 125
CC
CC RESULT 13
CC NRAM_IASH2 STANDARD; PRT; 468 AA.
CC AC P03478;
CC DT 21-JUL-1986 (Rel. 01, Created)
CC DT 01-FEB-1991 (Rel. 17, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Neuraminidase (EC 3.2.1.18).
CC GN NA.
CC OS Influenza A virus (strain A/Shearwater/Australia/72).
CC OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
CC OC Influenza A viruses; Influenzavirus A.
CC OX NCBI_TaxID=11461;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RX MEDLINE=89163262; PubMed=2922926;
CC RA Harley V.R., Ward C.W., Hudson P.J.;
CC RT "Molecular cloning and analysis of the N5 neuraminidase subtype from
CC an avian influenza virus.";
CC RL Virology 169:239-243(1989).
CC [2]
CC RP SEQUENCE OF 1-107 FROM N.A.
CC RN MEDLINE=83023028; PubMed=6896994;
CC RX

```

```

RA  Blok J., Air G.M.;
RT  "Variation in the membrane-insertion and 'stalk' sequences in eight
RL  subtypes of influenza type A virus neuraminidase.";
RL  Biochemistry 21:4001-4007(1982).
CC  -!- FUNCTION: Removes the terminal sialic acid from carbohydrate side
CC  chains of the host cell surface proteins and from the viral
CC  envelope. Such a reaction prevents self-aggregation and facilitate
CC  the mobility of the virus to and from the site of infection.
CC  -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2->3)-, alpha-(2->6)-,
CC  alpha-(2->8)-glycosidic linkages of terminal sialic residues in
CC  oligosaccharides, glycoproteins, glycolipids, colominic acid and
CC  synthetic substrates.
CC  -!- SUBUNIT: Homotetramer.
CC  -!- FUNCTION: Removes the terminal sialic acid from carbohydrate side
CC  chains of the host cell surface proteins and from the viral
CC  envelope. Such a reaction prevents self-aggregation and facilitate
CC  the mobility of the virus to and from the site of infection.
CC  -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2->3)-, alpha-(2->6)-,
CC  alpha-(2->8)-glycosidic linkages of terminal sialic residues in
CC  oligosaccharides, glycoproteins, glycolipids, colominic acid and
CC  synthetic substrates.
CC  -!- SUBUNIT: Homotetramer.
CC  -!- SUBCELLULAR LOCATION: VIRAL MEMBRANE. FORMS A MUSHROOM-SHAPED
CC  SPIKE ON THE SURFACE OF THE VIRION.
CC  -!- SIMILARITY: BELONGS TO FAMILY 34 OF GLYCOSYL HYDROLASES.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; M24740; AAA43672.1; -.
DR  EMBL; J02159; AAA43414.1; -.
DR  PIR; A00890; NMIVN5.
DR  HSP; P06820; 2BAT.
DR  InterPro; IPR001860; Glyco_hydro_34.
DR  Pfam; PF00064; neur; 1.
DR  ProDom; PD000431; Glyco_hydro_34; 1.
KW  Hydrolase; Glycosidase; Glycoprotein; Transmembrane.
FT  TRANSMEM 7 35 ANCHOR.
FT  DOMAIN 36 82 HYPERVARIABLE STALK REGION.
FT  DOMAIN 83 468 HEAD OF NEURAMINIDASE.
FT  ACT_SITE 267 267 PROBABLE.
FT  ACT_SITE 269 269 PROBABLE.
FT  CARBOHYD 46 46 N-LINKED (GLCNAC. .) (POTENTIAL).
FT  CARBOHYD 52 52 N-LINKED (GLCNAC. .) (POTENTIAL).
FT  CARBOHYD 65 65 N-LINKED (GLCNAC. .) (POTENTIAL).
FT  CARBOHYD 78 78 N-LINKED (GLCNAC. .) (POTENTIAL).
FT  CARBOHYD 138 138 N-LINKED (GLCNAC. .) (POTENTIAL).
FT  CARBOHYD 395 395 V -> G (IN REF. 2).
FT  CONFLICT 26 26 LGI -> WGT (IN REF. 2).
FT  CONFLICT 31 33 V -> A (IN REF. 2).
FT  CONFLICT 50 50 VRVET -> ARAEK (IN REF. 2).
FT  CONFLICT 55 59 YL -> HS (IN REF. 2).
FT  CONFLICT 69 70 C -> R (IN REF. 2).
FT  CONFLICT 84 84 KDNIGRIGSRGHIF -> NGHGRTRTGQEGTHS
FT  CONFLICT 94 107 (IN REF. 2).
SQ  SEQUENCE 468 AA; 51589 MW; 58A2335445B2D3F1F CRC64;

Query Match 67.4%; Score 29; DB 1; Length 468;
Best Local Similarity 55.6%; Pred. No. 61;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9
| : |||:|
Db 124 FFQTQALL 132

RESULT 14
NRAM_IACHI STANDARD; PRT; 470 AA.
AC P11485;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Neuraminidase (EC 3.2.1.18).
GN NA.
OS Influenza A virus (strain A/Chile/1/83).

```

```

OC  Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OX  Influenza A viruses; Influenzavirus A.
RN  NCBI_TaxID=11337;
RP  SEQUENCE FROM N.A.
RX  MEDLINE=88221459; PubMed=3369945;
RA  Schreier E., Roeske H., Driesel G., Kuenkel U., Petzold D.R.,
RA  Berlinghoff R., Michel S.;
RT  "Complete nucleotide sequence of the neuraminidase gene of the human
RT  influenza virus A/Chile/1/83 (H1N1). Brief report.";
RL  Arch. Virol. 99:271-276(1988).
CC  -!- FUNCTION: Removes the terminal sialic acid from carbohydrate side
CC  chains of the host cell surface proteins and from the viral
CC  envelope. Such a reaction prevents self-aggregation and facilitate
CC  the mobility of the virus to and from the site of infection.
CC  -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2->3)-, alpha-(2->6)-,
CC  alpha-(2->8)-glycosidic linkages of terminal sialic residues in
CC  oligosaccharides, glycoproteins, glycolipids, colominic acid and
CC  synthetic substrates.
CC  -!- SUBUNIT: Homotetramer.
CC  -!- SUBCELLULAR LOCATION: VIRAL MEMBRANE. FORMS A MUSHROOM-SHAPED
CC  SPIKE ON THE SURFACE OF THE VIRION.
CC  -!- SIMILARITY: BELONGS TO FAMILY 34 OF GLYCOSYL HYDROLASES.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; X15281; CAA33354.1; -.
DR  EMBL; M24783; AAA96671.1; ALT_SEQ.
DR  HSP; P03472; 2QWC.
DR  InterPro; IPR001860; Glyco_hydro_34.
DR  Pfam; PF00064; neur; 1.
DR  ProDom; PD000431; Glyco_hydro_34; 1.
KW  Hydrolase; Glycosidase; Glycoprotein; Transmembrane.
FT  TRANSMEM 7 35 ANCHOR.
FT  DOMAIN 36 90 HYPERVARIABLE STALK REGION.
FT  DOMAIN 91 470 HEAD OF NEURAMINIDASE.
FT  ACT_SITE 275 275 PROBABLE.
FT  ACT_SITE 277 277 PROBABLE.
FT  CARBOHYD 44 44 N-LINKED (GLCNAC. .) (POTENTIAL).
FT  CARBOHYD 58 58 N-LINKED (GLCNAC. .) (POTENTIAL).
FT  CARBOHYD 63 63 N-LINKED (GLCNAC. .) (POTENTIAL).
FT  CARBOHYD 70 70 N-LINKED (GLCNAC. .) (POTENTIAL).
FT  CARBOHYD 88 88 N-LINKED (GLCNAC. .) (POTENTIAL).
FT  CARBOHYD 146 146 N-LINKED (GLCNAC. .) (POTENTIAL).
FT  CARBOHYD 235 235 N-LINKED (GLCNAC. .) (POTENTIAL).
FT  CARBOHYD 365 365 N-LINKED (GLCNAC. .) (POTENTIAL).
FT  CARBOHYD 455 455 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ  SEQUENCE 470 AA; 51874 MW; 8647098A6C8CF25 CRC64;

Query Match 67.4%; Score 29; DB 1; Length 470;
Best Local Similarity 55.6%; Pred. No. 62;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9
| : |||:|
Db 132 FFQTQALL 140

RESULT 15
NRAM_IAKIE STANDARD; PRT; 470 AA.
AC P31348;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Neuraminidase (EC 3.2.1.18).
GN NA.

```

Tue Oct 7 15:52:37 2003

Search completed: October 7, 2003, 13:26:26  
Job time : 4.62025 secs

```
OS Influenza A virus (strain A/Kiev/59/79).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=11422;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=86077083; PubMed=4074401;
RX Beklemishev A.B., Blinov V.M., Vassilenko S.K., Golovin S.Y.,
RA Karginov V.A., Mamayev L.V., Netesov S.V., Petrov N.A.,
RA Safronov P.F.;
RT "Primary structure of a full-size DNA copy of the influenza virus
RT A/Kiev/59/79 (H1N1) neuraminidase gene.";
RL Bloorg. Khim. 11:1423-1426(1985).
CC -1- FUNCTION: Removes the terminal sialic acid from carbohydrate side
CC chains of the host cell surface proteins and from the viral
CC envelope. Such a reaction prevents self-aggregation and facilitate
CC the mobility of the virus to and from the site of infection.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2->3)-, alpha-(2->6)-,
CC alpha-(2->8)-glycosidic linkages of terminal sialic residues in
CC oligosaccharides, glycoproteins, glycolipids, colominic acid and
CC synthetic substrates.
CC -1- SUBUNIT: Homotrimer.
CC -1- SUBCELLULAR LOCATION: VIRAL MEMBRANE. FORMS A MUSHROOM-SHAPED
CC SPIKE ON THE SURFACE OF THE VIRION.
CC -1- SIMILARITY: BELONGS TO FAMILY 34 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M38335; AAA43435.1; -.
DR HSP; P03472; QWC.
DR InterPro; IPR001860; Glyco_hydro_34.
DR Pfam; PF00064; neur. 1.
DR ProDom; PD000431; Glyco_hydro_34; 1.
KW Hydrolase; Glycosidase; Glycoprotein; Transmembrane.
FT TRANSMEM 7 35
FT DOMAIN 36 91 ANCHOR
FT DOMAIN 92 470 HYPERVARIABLE STALK REGION.
FT ACT_SITE 275 275 PROBABLE.
FT ACT_SITE 277 277 PROBABLE.
FT DISULFID 92 417 BY SIMILARITY.
FT DISULFID 124 129 BY SIMILARITY.
FT DISULFID 184 231 BY SIMILARITY.
FT DISULFID 233 238 BY SIMILARITY.
FT DISULFID 279 292 BY SIMILARITY.
FT DISULFID 281 290 BY SIMILARITY.
FT DISULFID 318 335 BY SIMILARITY.
FT DISULFID 421 447 BY SIMILARITY.
FT CARBOHYD 44 44 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 58 58 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 63 63 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 68 68 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 88 88 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 146 146 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 235 235 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 285 285 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 365 365 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 455 455 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 470 AA; 51907 MW; 18A3917BFD2537D CRC64;
Query Match 67.4%; Score 29; DB 1; Length 470;
Best Local Similarity 55.6%; Pred. No. 62;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Qy 1 FLMAQGAML 9
Db 132 FFLTQGALL 140
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:18:33 ; Search time 15.038 Seconds  
(without alignments)  
154.441 Million cell updates/sec

Title: US-09-807-512-24  
Perfect score: 43  
Sequence: 1 FLMAQGAML 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues  
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : SPTREMBL\_23:\*
- 1: sp.archaea:\*
  - 2: sp.bacteria:\*
  - 3: sp.fungi:\*
  - 4: sp.human:\*
  - 5: sp.invertebrate:\*
  - 6: sp.mammal:\*
  - 7: sp.mhch:\*
  - 8: sp.organella:\*
  - 9: sp.phage:\*
  - 10: sp.plant:\*
  - 11: sp.rodent:\*
  - 12: sp.virus:\*
  - 13: sp.vertibrate:\*
  - 14: sp.unclassified:\*
  - 15: sp.rvirus:\*
  - 16: sp.bacteriap:\*
  - 17: sp.archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	100.0	58	095146	095146 homo sapien
2	43	100.0	109	095987	095987 homo sapien
3	34	79.1	192	085695	085695 streptomyce
4	34	79.1	1037	09AC25	Q9ac25 caulobacter
5	33	76.7	234	08GV55	Q8gv55 arabidopsis
6	33	76.7	236	08Y2U4	Q8y2u4 ralstonia s
7	32	74.4	156	08BUK5	Q8buk5 mus musculu
8	32	74.4	244	09H6F9	Q9h6f9 homo sapien
9	32	74.4	290	017572	017572 caenorhabdi
10	32	74.4	350	08XBI3	Q8xbi3 escherichia
11	32	74.4	350	08CXV2	Q8cxv2 escherichia
12	32	74.4	404	08K2W2	Q8k2w2 mus musculu
13	32	74.4	404	08JUC9	Q8juc9 mus musculu
14	32	74.4	404	09D2D0	Q9d2d0 mus musculu
15	32	74.4	406	09BSE4	Q9bse4 homo sapien
16	32	74.4	469	08B198	Q8b198 influenza a

17	32	74.4	469	12	Q8B197	Q8b197 influenza a
18	32	74.4	904	16	Q8YQ47	Q8yq47 anabaena sp
19	32	74.4	1101	4	Q9NXY9	Q9nxy9 homo sapien
20	32	74.4	3898	12	O11994	O11994 mucosal dis
21	32	74.4	3907	12	Q96662	Q96662 mucosal dis
22	32	74.4	3975	12	Q65815	Q65815 mucosal dis
23	32	74.4	4983	12	O11993	O11993 mucosal dis
24	31	72.1	222	5	Q9VIT6	Q9vit6 drosophila
25	31	72.1	476	10	Q9SIA4	Q9sia4 arabidopsis
26	31	72.1	476	10	Q9SIA3	Q9sia3 arabidopsis
27	31	72.1	557	16	Q8F5R3	Q8f5r3 leptospira
28	31	72.1	595	16	Q8D902	Q8d902 vibrio vuln
29	31	72.1	815	10	O22113	O22113 chlorococcu
30	31	72.1	3343	12	Q91HK6	Q91hk6 papaya ring
31	31	72.1	3344	12	P90213	P90213 papaya ring
32	30	69.8	111	10	O23995	O23995 hordeum vul
33	30	69.8	269	5	Q9VIT5	Q9vit5 drosophila
34	30	69.8	302	4	Q9HA01	Q9ha01 homo sapien
35	30	69.8	380	4	Q96ER9	Q96er9 homo sapien
36	30	69.8	382	5	Q8I8I5	Q8i8i5 mastigamoeb
37	30	69.8	389	17	Q8TQ40	Q8tq40 methanosarc
38	30	69.8	389	17	Q8PTR4	Q8ptr4 methanosarc
39	30	69.8	415	16	Q8YRJ4	Q8yrj4 anabaena sp
40	30	69.8	444	10	O22402	O22402 nicotiana t
41	30	69.8	455	16	Q8E8P4	Q8e8p4 shewanella
42	30	69.8	482	16	Q8XU27	Q8xu27 ralstonia s
43	30	69.8	492	2	Q9AHT7	Q9aht7 streptococc
44	30	69.8	501	2	Q9Z4J8	Q9z4j8 streptococc
45	30	69.8	501	16	Q8DP07	Q8dp07 streptococc

ALIGNMENTS

RESULT 1

ID O95146 PRELIMINARY; PRT; 58 AA.  
AC O95146:  
DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE LAGE-2ALT protein isoform.  
GN LAGE-2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Wang R.-F., Johnston S.L., Zeng G., Topalian S.L.,  
RA Schwartzentruber D.J., Rosenberg S.A.;  
RT "A breast and Melanoma-Shared Tumor Antigen: T Cell Responses to  
RT Antigenic Peptides Translated from Different Open Reading Frames.";  
RT J. Immunol. 161:3596-3606(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Lethe B.G.;  
RN Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RX MEDLINE=99454989; PubMed=10523621;  
RA De Smet C., Lurquin C., Lethe B., Martelange V., Boon T.;  
RT "DNA methylation is the primary silencing mechanism for a set of germ  
RT line- and tumor-specific genes with a CpG-rich promoter.";  
DR EMBL; AF038567; AAD05203.1; -;  
DR EMBL; AJ275977; CAB76944.1; -;  
SQ SEQUENCE 58 AA; 6188 MW; ED12057564BC7EF2 CRC64;

Query Match 100.0%; Score 43; DB 4; Length 58;  
Best Local Similarity 100.0%; Pred. No. 0.068;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RA	Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J., Kinashi H., Hopwood D.A.; "A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome."; Mol. Microbiol. 21:77-96(1996).
RT	[5]
RL	SEQUENCE FROM N.A.
RN	SPECIES-S.coelicolor; STRAIN-A3(2) / M145;
RP	MEDLINE=21996410; PubMed=12000953;
RC	Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L., Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H., Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M., Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S., Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S., Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S., Seeger K., Saunders D., Sharp S., Squares R., Taylor K., Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J., Hopwood D.A.;
RD	"Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2)."; Nature 417:141-147(2002).
RE	-!- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL REGULATORS.
RF	EMBL; AF072709; AAC25764.1; -.
RG	EMBL; AL939106; CAC14337.1; -.
RH	InterPro: IPR001647; HTH_Tetr.
RI	Pfam: PF00440; tetr; 1.
RJ	PRINTS: PR00445; HTHHTETR.
RK	DNA-binding; transcription regulation; Complete proteome.
RL	SEQUENCE 192 AA; 20556 MW; 7DBAA07EBBA8DDFC CRC64;
RM	Query Match 79.1%; Score 34; DB 16; Length 192;
RN	Best Local Similarity 66.7%; Pred. No. 21;
RO	Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
RP	QY 1 FLMAQGAML 9 II::III:
RQ	Db 150 FLAEGAMV 158
RS	RESULT 4
RT	O9AC25 PRELIMINARY; PRT; 1037 AA.
RU	ID O9AC25
RV	AC O9AC25
RW	DT 01-JUN-2001 (TReMBLrel. 17, Created)
RX	DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
RY	DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
RZ	DE Putative translational initiation factor IF-2.
SA	GN CC0042.
SB	OS Caulobacter crescentus.
SC	OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
SD	OC Caulobacteraceae; Caulobacter.
SE	OX NCBI_TaxID=155892;
SF	[1]
SG	RP SEQUENCE FROM N.A.
SH	RC STRAIN=ATCC 19089 / CB15;
SI	MEADLINE=21173698; PubMed=11259647;
SJ	RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E., Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R., Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B., DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H., Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K., O., Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O., Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
SK	RA "Complete genome sequence of Caulobacter crescentus."; Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
SL	RT -!- FUNCTION: IF-2, ONE OF THE ESSENTIAL COMPONENTS FOR THE INITIATION OF PROTEIN SYNTHESIS IN VITRO, PROTECTS FORMYL METHIONYL-TRNA FROM SPONTANEOUS HYDROLYSIS AND PROMOTES ITS BINDING TO THE 30S RIBOSOMAL SUBUNIT. IT IS ALSO INVOLVED IN THE HYDROLYSIS OF GTP DURING THE FORMATION OF THE 70S RIBOSOMAL COMPLEX (BY SIMILARITY).
SM	-!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
SN	-!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.

CC -!- SIMILARITY: BELONGS TO THE IF-2 FAMILY.  
 DR EMBL; AE005678; AAK22030.1; -.  
 DR TIGR; CC0042; -.

DR InterPro; IPR004161; EFTULD2.  
 DR InterPro; IPR000795; EF\_GTPbind.  
 DR InterPro; IPR000178; IF2.  
 DR InterPro; IPR006847; IF2\_N.  
 DR InterPro; IPR002965; P\_Rich\_extensn.  
 DR InterPro; IPR005225; Small\_GTP.  
 DR Pfam; PF00009; GTP\_EFTU; 1.  
 DR Pfam; PF03144; GTP\_EFTULD2; 2.  
 DR Pfam; PF04760; IF2\_N; 1.  
 DR PRINTS; PR01217; PRICHEXTENS.  
 DR PRODOM; PD186100; IF2; 1.  
 DR TIGRFAMS; TIGR00487; IF-2; 1.  
 DR TIGRFAMS; TIGR00231; small\_GTP; 1.  
 DR PROSITE; PS01176; IF2; 1.  
 KW GTP-binding; Initiation factor; Protein biosynthesis;  
 KW Complete proteome.  
 SQ SEQUENCE 1037 AA; 110678 MW; 5759830727DC5716 CRC64;

Query Match 79.1%; Score 34; DB 16; Length 1037;  
 Best Local Similarity 77.8%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9  
 ||| |||  
 DB 479 FLMRQGVML 487

RESULT 5  
 Q8GV55 PRELIMINARY; PRT; 234 AA.  
 AC Q8GV55;  
 DT 01-MAR-2003 (TRENBLrel. 23, Created)  
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
 DE Leafy cotyledon 1-like LIL protein.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kwong R.W., Bui A.Q., Lee H., Kwong L.W., Fischer R.L., Goldberg R.B.,  
 RA Harada J.J.;  
 RT "Leafy Cotyledon 1-Like is an Essential Regulator of Arabidopsis  
 RT Embryo Development."  
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY138461; AAN15924.1; -.  
 SQ SEQUENCE 234 AA; 26136 MW; EC100E830E5CBB31 CRC64;

Query Match 76.7%; Score 33; DB 10; Length 234;  
 Best Local Similarity 75.0%; Pred. No. 42;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAQGAM 8  
 ||||:|  
 DB 28 FLMAEGSM 35

RESULT 6  
 Q8Y2U4 PRELIMINARY; PRT; 236 AA.  
 AC Q8Y2U4;  
 DT 01-MAR-2002 (TRENBLrel. 20, Created)  
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)  
 DE Putative permease transmembrane protein.  
 GN RSC0238 OR RS00670.  
 OS Ralstonia solanacearum (Pseudomonas solanacearum).  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;

OC Ralstoniaceae; Ralstonia.  
 OX NCBI\_TaxID=305;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GM11000;  
 RX MEDLINE=21681879; PubMed=11823852;  
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,  
 RA Arlat M., Billault A., Bottier P., Camus J.C., Cattolico L.,  
 RA Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,  
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,  
 RA Sigulier P., Thebault P., Whalen M., Wincker P., Levy M.,  
 RA Weissenbach J., Boucher C.A.;  
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum."  
 RL Nature 415:497-502(2002).  
 DR EMBL; AL646058; CAD13766.1; -.  
 KW Complete proteome.  
 SQ SEQUENCE 236 AA; 24140 MW; 64417628C19BE432 CRC64;

Query Match 76.7%; Score 33; DB 16; Length 236;  
 Best Local Similarity 66.7%; Pred. No. 42;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9  
 ||:|:|  
 DB 19 FLTAQGSLL 27

RESULT 7  
 Q8BUK5 PRELIMINARY; PRT; 156 AA.  
 AC Q8BUK5;  
 DT 01-MAR-2003 (TRENBLrel. 23, Created)  
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)  
 DE Putative ankyrin-repeat containing protein homolog.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Heart;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium,  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs."  
 RL Nature 420:563-573(2002).  
 DR EMBL; AK084601; BAC39224.1; -.  
 SQ SEQUENCE 156 AA; 16591 MW; 5770AB021A51B810 CRC64;

Query Match 74.4%; Score 32; DB 11; Length 156;  
 Best Local Similarity 66.7%; Pred. No. 46;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9  
 ||:|:|:|  
 DB 24 FLTAQGASI 32

RESULT 8  
 Q9H6F9 PRELIMINARY; PRT; 244 AA.  
 AC Q9H6F9;  
 DT 01-MAR-2001 (TRENBLrel. 16, Created)  
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)  
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)  
 DE Hypothetical protein FLJ22313.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]

RP SEQUENCE FROM N.A.  
RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,  
RA Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T.,  
RA Nakamura Y., Iwagaki T., Sugano S.,  
RT "NEDO human cDNA sequencing project."  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK025966; BAB15300.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 244 AA; 27270 MW; CBF239890489B918 CRC64;  
  
Query Match 74.4%; Score 32; DB 4; Length 244;  
Best Local Similarity 66.7%; Pred. No. 72;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 1 FLMAQGAML 9  
Db 140 FIMVGMAML 148  
I:| | | | |  
140 FIMVGMAML 148  
  
RESULT 9  
O17572 PRELIMINARY; PRT; 290 AA.  
ID O17572  
AC O17572  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE C06B8.6 protein.  
GN C06B8.6  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Steward C.A.;  
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99069613; PubMed=9851916;  
RA none;  
RT "Genome sequence of the nematode C.elegans: A platform for  
investigating biology."  
RL Science 282:2012-2018(1998).  
DR EMBL; Z81463; CAB03849.1; -  
DR WormPep; C06B8.6; CE07974.  
DR InterPro; IPR003003; 7TM\_chemo2.  
DR InterPro; IPR000168; 7TM\_nematode.  
DR Pfam; PF01604; 7tm\_5; 1.  
SQ SEQUENCE 290 AA; 33363 MW; 0B4900E551FA3CD5 CRC64;  
  
Query Match 74.4%; Score 32; DB 5; Length 290;  
Best Local Similarity 77.8%; Pred. No. 85;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 1 FLMAQGAML 9  
Db 158 FLCAQGAFL 166  
I:| | | | |  
158 FLCAQGAFL 166  
  
RESULT 10  
Q8XB13 PRELIMINARY; PRT; 350 AA.  
ID Q8XB13  
AC Q8XB13  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Putative ARAC-type regulatory protein.  
GN YFEG OR Z3702 OR ECS3308.  
OS Escherichia coli O157:H7.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=83334;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
RX MEDLINE=21074935; PubMed=11206551;  
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
RA Postai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamocousis K.,  
RA Grodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
RA Welch R.A., Blattner F.R.;  
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";  
RL Nature 409:529-533(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O157:H7 / RIMD 0509952;  
RX MEDLINE=21156231; PubMed=11258796;  
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
RA Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,  
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;  
RT "Complete genome sequence of enterohemorrhagic Escherichia coli  
O157:H7 and genomic comparison with a laboratory strain K-12";  
RL DNA Res. 8:11-22(2001).  
DR EMBL; AE005474; AAG57555.1; -  
DR EMBL; AP002561; BAB36731.1; -  
DR InterPro; IPR000005; HTHARAC.  
DR Pfam; PF00165; HTH\_ARAC; 2.  
DR SMART; SM00342; HTH\_ARAC; 1.  
DR PROSITE; PS00041; HTH\_ARAC\_FAMILY\_1; 1.  
DR PROSITE; PS01124; HTH\_ARAC\_FAMILY\_2; 1.  
KW Complete proteome.  
SQ SEQUENCE 350 AA; 40086 MW; 1F9E97EC2E00E1D5 CRC64;  
  
Query Match 74.4%; Score 32; DB 16; Length 350;  
Best Local Similarity 87.5%; Pred. No. 1e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 2 LMAQGAML 9  
Db 214 LMANGAML 221  
I:| | | | |  
214 LMANGAML 221  
  
RESULT 11  
Q8CX2 PRELIMINARY; PRT; 350 AA.  
ID Q8CX2  
AC Q8CX2;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Ethanolamine operon regulatory protein.  
GN YFEG OR C2971.  
OS Escherichia coli O6.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=217992;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O6:H1 / CFT073 / ATCC 700928;  
RX MEDLINE=22388234; PubMed=12471157;  
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,  
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,  
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,  
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;  
RT "Extensive mosaic structure revealed by the complete genome sequence  
of uropathogenic Escherichia coli";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).  
DR EMBL; AE016764; AAN81421.1; -  
KW Complete proteome.  
SQ SEQUENCE 350 AA; 40119 MW; 0BCE98BC3500E1D5 CRC64;  
  
Query Match 74.4%; Score 32; DB 16; Length 350;  
Best Local Similarity 87.5%; Pred. No. 1e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;



```

QY 2 LMAQGAML 9
Db 214 LMAMGAML 221

RESULT 12
Q8K2W2 PRELIMINARY; PRT; 404 AA.
AC Q8K2W2;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE RIKEN cDNA 5031400M07 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC029691; AAH29691.1; -.
DR InterPro: IPR000626; Ubiquitin.
DR PROSITE: PS50053; UBIQUITIN_2; 1.
SQ SEQUENCE 404 AA; 44531 MW; EED6273F7300BF13 CRC64;

Query Match 74.4%; Score 32; DB 11; Length 404;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9
Db 301 FIMVGMAML 309
I:| | | | |

RESULT 13
Q9JJC9 PRELIMINARY; PRT; 404 AA.
AC Q9JJC9;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Brain cDNA, clone MNCB-2040.
GN 5031400M07RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Osada N., Kusuda J., Tanuma R., Ito A., Hirata M., Sugano S.,
RA Hashimoto K.;
RT "isolation of full-length cDNA clones from mouse brain cDNA library
made by oligo-capping method.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB041580; BAA95064.1; -.
DR MGD; MGI:1915393; 5031400M07RIK.
DR InterPro: IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
SQ SEQUENCE 404 AA; 44501 MW; EECBEEEB7ADD76C7 CRC64;

Query Match 74.4%; Score 32; DB 11; Length 404;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9
Db 301 FIMVGMAML 309
I:| | | | |

RESULT 14
Q9D2D0 PRELIMINARY; PRT; 404 AA.
AC Q9D2D0;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE 5031400M07RIK protein.
GN 5031400M07RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary, and Uterus;
RX MEDLINE=21085560; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK019862; BAB31889.1; -.
DR MGD; MGI:1915393; 5031400M07RIK.
DR InterPro: IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
SQ SEQUENCE 404 AA; 44515 MW; 9879095F2BDD76C3 CRC64;

Query Match 74.4%; Score 32; DB 11; Length 404;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9
Db 301 FIMVGMAML 309
I:| | | | |

RESULT 15
Q9BSE4 PRELIMINARY; PRT; 406 AA.
AC Q9BSE4;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Similar to RIKEN cDNA 5031400M07 gene (Similar to homocysteine-
inducible protein).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]

```

Tue Oct 7 15:52:38 2003

RP SEQUENCE FROM N.A.  
RC TISSUE-Colon:  
RA Strausberg R.;  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC005091; AAH05091.1; -;  
DR EMBL; BC020264; AAH20264.1; -;  
DR InterPro; IPR000626; Ubiquitin.  
DR Pfam; PF00240; ubiquitin; 1.  
DR PROSITE; PS50053; UBIQUITIN\_2; 1.  
SQ SEQUENCE 406 AA; 45177 MW; E2147DA53E18CE3A CRC64;

Query Match 74.4%; Score 32; DB 4; Length 406;  
Best Local Similarity 66.7%; Pred. No. 1.2e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9  
|:| | | |  
Db 302 FIMVYGAML 310

Search completed: October 7, 2003, 13:31:01  
Job time : 17.038 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:21:07 ; Search time 4.89873 Seconds  
(without alignments)  
77.734 Million cell updates/sec

Title: US-09-807-512-24  
Perfect score: 43  
Sequence: 1 FLMAQGAML 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A-COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B-COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A-COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B-COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	32	74.4	2616	6 5206163-3	Patent No. 5206163
2	31	72.1	592	2 US-08-366-490-2	Sequence 2, Appli
3	31	72.1	592	2 US-08-366-490-6	Sequence 6, Appli
4	31	72.1	592	3 US-08-860-483A-2	Sequence 2, Appli
5	31	72.1	592	3 US-08-860-483A-6	Sequence 6, Appli
6	31	72.1	596	3 US-08-860-483A-12	Sequence 12, Appl
7	31	72.1	596	3 US-08-860-483A-13	Sequence 13, Appl
8	31	72.1	632	2 US-08-366-490-8	Sequence 8, Appli
9	30	69.8	68	4 US-09-328-352-991A-22345	Sequence 22345, A
10	30	69.8	445	4 US-09-328-352-991A-22345	Sequence 6669, Ap
11	30	69.8	471	3 US-08-961-083-158	Sequence 158, App
12	30	69.8	471	4 US-09-536-784-158	Sequence 158, App
13	30	69.8	1068	4 US-09-215-694-7	Sequence 7, Appli
14	29	67.4	62	4 US-09-615-192A-300	Sequence 300, App
15	29	67.4	462	4 US-09-252-991A-21817	Sequence 21817, A
16	29	67.4	469	3 US-08-686-968C-229	Sequence 229, App
17	28	65.1	333	4 US-09-444-336-10	Sequence 10, Appl
18	28	65.1	351	4 US-09-186-276B-23	Sequence 23, Appl
19	28	65.1	351	4 US-08-842-445-23	Sequence 23, Appl
20	28	65.1	351	4 US-09-186-188B-23	Sequence 23, Appl
21	28	65.1	353	4 US-09-186-276B-32	Sequence 32, Appl
22	28	65.1	353	4 US-08-842-445-32	Sequence 32, Appl
23	28	65.1	353	4 US-09-186-188B-32	Sequence 32, Appl
24	28	65.1	405	4 US-09-252-991A-29407	Sequence 29407, A
25	28	65.1	495	4 US-09-328-352-7210	Sequence 7210, Ap
26	28	65.1	512	4 US-09-134-001C-4349	Sequence 4349, Ap
27	28	65.1	613	4 US-09-328-352-5116	Sequence 5116, Ap

28	65.1	3798	3	US-09-335-409-6	Sequence 6, Appli
29	65.1	3798	4	US-09-568-102-6	Sequence 6, Appli
30	65.1	3798	4	US-09-567-969-6	Sequence 6, Appli
31	65.1	3798	4	US-09-568-480-6	Sequence 6, Appli
32	65.1	3798	4	US-09-568-486-6	Sequence 6, Appli
33	65.1	3798	4	US-09-568-472-6	Sequence 6, Appli
34	65.1	3798	4	US-09-567-899-6	Sequence 6, Appli
35	65.1	3798	4	US-08-152-721B-18	Sequence 18, Appli
36	62.8	178	4	US-09-252-991A-29942	Sequence 29942, A
37	62.8	179	4	US-09-615-192A-299	Sequence 299, App
38	62.8	185	4	US-09-252-991A-25713	Sequence 25713, A
39	62.8	193	4	US-09-328-352-4173	Sequence 4173, Ap
40	62.8	274	2	US-08-701-191A-33	Sequence 33, Appl
41	62.8	287	4	US-09-134-001C-4454	Sequence 4454, Ap
42	62.8	293	4	US-09-252-991A-28184	Sequence 28184, A
43	62.8	320	4	US-09-252-991A-30322	Sequence 30322, A
44	62.8	332	2	US-08-637-763B-6	Sequence 6, Appli
45	62.8	332	3	US-09-170-354-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1  
5206163-3  
; Patent No. 5206163  
; APPLICANT: RENARD, ANDRE-DINA, DINO; MARTIAL, JOSEPH  
; TITLE OF INVENTION: DNA ENCODING BOVINE DIARRHEA  
; VIRUS PROTEIN  
; NUMBER OF SEQUENCES: 3  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/550,816  
; FILING DATE: 06-JUL-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 331,037  
; FILING DATE: 29-MAR-1989  
; APPLICATION NUMBER: 752,981  
; FILING DATE: 08-JUL-1985  
; SEQ ID NO:3:  
; LENGTH: 2616  
5206163-3

Query Match 74.4%; Score 32; DB 6; Length 2616;  
Best Local Similarity 87.5%; Pred. No. 2.7e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	2	LMAQGAML 9
DB	1775	LMAQGRML 1782

RESULT 2  
US-08-366-490-2  
; Sequence 2, Application US/08366490  
; Patent No. 5877403  
; GENERAL INFORMATION:  
; APPLICANT: McMaster, J. Russell  
; APPLICANT: Boeshore, Maury L  
; APPLICANT: Triccoli, David M  
; APPLICANT: Reynolds, John F  
; APPLICANT: Carney, Kim J  
; TITLE OF INVENTION: PAPAYA RINGSPOT VIRUS PROTEASE GENE  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fitzpatrick, Cella, Harper, and Scinto  
; STREET: 277 Park Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10172-0194  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC Compatible

Matches	7;	Conservative	0;	Mismatches	2;	Indels	0;	Gaps	0;
QY	1	FLMAQGAML 9							
Db	95	FTWAGGAML 103							
<p>RESULT 4</p> <p>US-08-860-483A-2</p> <p>; Sequence 2, Application US/08860483A</p> <p>; Patent No. 6046384</p> <p>; GENERAL INFORMATION:</p> <p>; APPLICANT: McMaster, J. R. R.</p> <p>; APPLICANT: Boeshore, Maury L.</p> <p>; APPLICANT: Tricoli, David M.</p> <p>; APPLICANT: Reynolds, John F.</p> <p>; APPLICANT: Carney, Kim J.</p> <p>; APPLICANT: Slighton, Jerry L.</p> <p>; APPLICANT: Gonsalves, Dennis</p> <p>; TITLE OF INVENTION: Papaya Ringspot Virus Nta Protease Gene</p> <p>; NUMBER OF SEQUENCES: 13</p> <p>; CORRESPONDENCE ADDRESS:</p> <p>; ADDRESSEE: Rocky, Milnamow &amp; Katz</p> <p>; STREET: 180 N. Stetson Avenue, 2 Prudential Plaza,</p> <p>; CITY: Suite 4700</p> <p>; STATE: IL</p> <p>; COUNTRY: USA</p> <p>; ZIP: 60601</p> <p>; COMPUTER READABLE FORM:</p> <p>; MEDIUM TYPE: Floppy disk</p> <p>; COMPUTER: IBM PC compatible</p> <p>; OPERATING SYSTEM: PC-DOS/MS-DOS</p> <p>; SOFTWARE: PatentIn Release #1.0, Version #1.30</p> <p>; CURRENT APPLICATION DATA:</p> <p>; APPLICATION NUMBER: US/08/860.483A</p> <p>; FILING DATE: 26-JUN-1997</p> <p>; CLASSIFICATION: 800</p> <p>; ATTORNEY/AGENT INFORMATION:</p> <p>; NAME: Mueller, Lisa V.</p> <p>; REGISTRATION NUMBER: 38,978</p> <p>; REFERENCE/DOCKET NUMBER: SVS3801P0091US</p> <p>; TELECOMMUNICATION INFORMATION:</p> <p>; TELEPHONE: 3126165400</p> <p>; TELEFAX: 3126165460</p> <p>; INFORMATION FOR SEQ ID NO: 2:</p> <p>; SEQUENCE CHARACTERISTICS:</p> <p>; LENGTH: 592 amino acids</p> <p>; TYPE: amino acid</p> <p>; TOPOLOGY: linear</p> <p>; MOLECULE TYPE: protein</p> <p>US-08-860-483A-2</p> <p>Query Match 72.1%; Score 31; DB 3; Length 592;</p> <p>Best Local Similarity 77.8%; Pred. No. 93;</p> <p>Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;</p>									
QY	1	FLMAQGAML 9							
Db	95	FTWAGGAML 103							
<p>RESULT 5</p> <p>US-08-860-483A-6</p> <p>; Sequence 6, Application US/08860483A</p> <p>; Patent No. 6046384</p> <p>; GENERAL INFORMATION:</p> <p>; APPLICANT: McMaster, J. R. R.</p> <p>; APPLICANT: Boeshore, Maury L.</p> <p>; APPLICANT: Tricoli, David M.</p> <p>; APPLICANT: Reynolds, John F.</p> <p>; APPLICANT: Carney, Kim J.</p> <p>; APPLICANT: Slighton, Jerry L.</p>									

APPLICANT: Gonsalves, Dennis  
TITLE OF INVENTION: Papaya Ringspot Virus N1a Protease Gene  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rocky, Milnamow & Katz  
STREET: 180 N. Stetson Avenue, 2 Prudential Plaza,  
CITY: Suite 4700  
STATE: IL  
COUNTRY: USA  
ZIP: 60601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/860,483A  
FILING DATE: 26-JUN-1997  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Mueller, Lisa V.  
REGISTRATION NUMBER: 38,978  
REFERENCE/DOCKET NUMBER: SVS3801P0091US  
TELEPHONE: 3126165400  
TELEFAX: 3126165460  
INFORMATION FOR SEQ ID NO: 6:  
LENGTH: 592 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-860-483A-6

Query Match 72.1%; Score 31; DB 3; Length 592;  
Best Local Similarity 77.8%; Pred. No. 93;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9  
Db 95 FTMAGGAML 103

RESULT 6  
US-08-860-483A-12  
Sequence 12, Application US/08860483A  
Patent No. 6046384  
GENERAL INFORMATION:  
APPLICANT: McMaster, J. R.  
APPLICANT: Boeshore, Maury L.  
APPLICANT: Tricoli, David M.  
APPLICANT: Reynolds, John F.  
APPLICANT: Carney, Kim J.  
APPLICANT: Slighton, Jerry L.  
APPLICANT: Gonsalves, Dennis  
TITLE OF INVENTION: Papaya Ringspot Virus N1a Protease Gene  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rocky, Milnamow & Katz  
STREET: 180 N. Stetson Avenue, 2 Prudential Plaza,  
CITY: Suite 4700  
STATE: IL  
COUNTRY: USA  
ZIP: 60601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/860,483A

FILING DATE: 26-JUN-1997  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Mueller, Lisa V.  
REGISTRATION NUMBER: 38,978  
REFERENCE/DOCKET NUMBER: SVS3801P0091US  
TELEPHONE: 3126165400  
TELEFAX: 3126165460  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 596 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: protein  
US-08-860-483A-12

Query Match 72.1%; Score 31; DB 3; Length 596;  
Best Local Similarity 77.8%; Pred. No. 93;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9  
Db 95 FTMAGGAML 103

RESULT 7  
US-08-860-483A-13  
Sequence 13, Application US/08860483A  
Patent No. 6046384  
GENERAL INFORMATION:  
APPLICANT: McMaster, J. R.  
APPLICANT: Boeshore, Maury L.  
APPLICANT: Tricoli, David M.  
APPLICANT: Reynolds, John F.  
APPLICANT: Carney, Kim J.  
APPLICANT: Slighton, Jerry L.  
APPLICANT: Gonsalves, Dennis  
TITLE OF INVENTION: Papaya Ringspot Virus N1a Protease Gene  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rocky, Milnamow & Katz  
STREET: 180 N. Stetson Avenue, 2 Prudential Plaza,  
CITY: Suite 4700  
STATE: IL  
COUNTRY: USA  
ZIP: 60601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/860,483A  
FILING DATE: 26-JUN-1997  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Mueller, Lisa V.  
REGISTRATION NUMBER: 38,978  
REFERENCE/DOCKET NUMBER: SVS3801P0091US  
TELEPHONE: 3126165400  
TELEFAX: 3126165460  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 596 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: protein  
US-08-860-483A-13

Tue Oct 7 15:52:35 2003

```
Query Match      72.1%; Score 31; DB 3; Length 596;
Best Local Similarity 77.8%; Pred. No. 93;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 FLMAQGAML 9
Db      95 FTWAGGAML 103

RESULT 8
US-08-366-490-8
; Sequence 8, Application US/08366490
; Patent No. 5877403
; GENERAL INFORMATION:
; APPLICANT: McMaster, J. Russell
; APPLICANT: Boeshore, Maury L
; APPLICANT: Tricoll, David M
; APPLICANT: Reynolds, John F
; APPLICANT: Carney, Kim J
; TITLE OF INVENTION: PAPAYA RINGSPOT VIRUS PROTEASE GENE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Fitzpatrick, Cella, Harper, and Scinto
; STREET: 277 Park Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10172-0194
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/366.490
; FILING DATE: 30-DEC-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitzpatrick, Cella, Harper, and Scinto
; REFERENCE/DOCKET NUMBER: 4869
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-2400
; TELEFAX: 212-758-2982
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 632 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-366-490-8

Query Match      72.1%; Score 31; DB 2; Length 632;
Best Local Similarity 77.8%; Pred. No. 99;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 FLMAQGAML 9
Db      99 FTWAGGAML 107

RESULT 9
US-09-252-991A-22345
; Sequence 22345, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788

Query Match      69.8%; Score 30; DB 4; Length 68;
Best Local Similarity 85.7%; Pred. No. 15;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      3 MAQGAML 9
Db      17 LAQGAML 23

US-09-252-991A-22345
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22345
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-22345

US-09-328-352-6669
; Sequence 6669, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328.352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6669
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
; US-09-328-352-6669

Query Match      69.8%; Score 30; DB 4; Length 445;
Best Local Similarity 66.7%; Pred. No. 11e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 FLMAQGAML 9
Db      50 YLLALGAML 58

RESULT 11
US-08-961-083-158
; Sequence 158, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961.083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER:
; FILING DATE:
```

ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB340P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 158:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 471 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-961-083-158

Query Match 69.8%; Score 30; DB 3; Length 471;  
Best Local Similarity 66.7%; Pred. No. 1.2e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9  
|| |||||  
Db 151 FLNGOGAMM 159

## RESULT 12

US-09-536-784-158  
Sequence 158, Application US/09536784  
Patent No. 6573082

## GENERAL INFORMATION:

APPLICANT: Choi et. al.  
TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines  
NUMBER OF SEQUENCES: 452  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2

## SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/536,784  
FILING DATE: 30-Oct-1997

CLASSIFICATION: <Unknown>  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 08/961,083  
FILING DATE: OCT-30-1997

## ATTORNEY/AGENT INFORMATION:

NAME: Michelle S. Marks  
REGISTRATION NUMBER: 41,971  
REFERENCE/DOCKET NUMBER: PB340P3

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512

## INFORMATION FOR SEQ ID NO: 158:

## SEQUENCE CHARACTERISTICS:

LENGTH: 471 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

## MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 158:

Query Match 69.8%; Score 30; DB 4; Length 471;  
Best Local Similarity 66.7%; Pred. No. 1.2e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

US-09-536-784-158

QY 1 FLMAQGAML 9  
|| |||||  
Db 151 FLNGOGAMM 159

## RESULT 13

US-09-215-694-7  
Sequence 7, Application US/09215694B  
Patent No. 6391583

## GENERAL INFORMATION:

APPLICANT: Wisconsin Alumni Research Foundation  
APPLICANT: Hutchinson, Charles R.  
APPLICANT: Kennedy, Jonathan n.m.i.  
APPLICANT: Park, Cheonseek n.m.i.  
TITLE OF INVENTION: METHOD OF PRODUCING ANTIHYPERCHOLESTEROLEMIC AGENTS  
FILE REFERENCE: 960296.95718  
CURRENT APPLICATION NUMBER: US/09/215,694B  
CURRENT FILING DATE: 1999-12-18  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 7  
LENGTH: 1068  
TYPE: PRT  
ORGANISM: Aspergillus terreus  
US-09-215-694-7

Query Match 69.8%; Score 30; DB 4; Length 1068;

Best Local Similarity 87.5%; Pred. No. 2.8e+02;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LMAQGAML 9  
|| |||||  
Db 954 LEAQGAML 961

## RESULT 14

US-09-615-192A-300  
Sequence 300, Application US/09615192A  
Patent No. 6410718

## GENERAL INFORMATION:

APPLICANT: Bloksberg, Leonard N.  
APPLICANT: Havukkala, Ilkka  
TITLE OF INVENTION: Materials and Methods for the  
MODIFICATION OF PLANT LIGNIN CONTENT  
FILE REFERENCE: 11000.1003c4U  
CURRENT APPLICATION NUMBER: US/09/615,192A  
CURRENT FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 08/975,316  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: US 08/713,000  
PRIOR FILING DATE: 1996-09-11  
PRIOR APPLICATION NUMBER: US 09/169,789  
PRIOR FILING DATE: 1998-10-09  
NUMBER OF SEQ ID NOS: 405  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 300  
LENGTH: 62  
TYPE: PRT  
ORGANISM: Eucalyptus grandis  
US-09-615-192A-300

Query Match 67.4%; Score 29; DB 4; Length 62;

Best Local Similarity 66.7%; Pred. No. 23;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9  
|| |||||  
Db 18 FCMAQGEVL 26

## RESULT 15

US-09-252-991A-21817

Sequence 21817, Application US/09252991A

Tue Oct 7 15:52:35 2003

; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 21817  
; LENGTH: 462  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-21817

Query Match 67.4%; Score 29; DB 4; Length 462;  
Best Local Similarity 55.6%; Pred. No. 1.9e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9  
;| | | | |  
Db 221 YLFAQGAVI 229

Search completed: October 7, 2003, 13:34:06  
Job time : 5.89873 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:31:20 ; Search time 9.79747 Seconds  
(without alignments)  
145.335 Million cell updates/sec

Title: US-09-807-512-24  
Perfect score: 43  
Sequence: 1 FLMAQGAML 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues  
Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PTCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PTCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32	74.4	1101	15 US-10-177-293-106	Sequence 106, App
2	31	72.1	88	9 US-09-925-301-1327	Sequence 1327, Ap
3	31	72.1	88	15 US-10-106-698-4877	Sequence 4877, Ap
4	31	72.1	184	15 US-10-156-761-14999	Sequence 14999, A
5	30	69.8	402	12 US-10-181-590-8	Sequence 8, Appli
6	30	69.8	471	9 US-09-765-272-158	Sequence 158, App
7	30	69.8	490	12 US-09-769-744A-98	Sequence 98, Appl
8	30	69.8	1011	12 US-10-032-585-7548	Sequence 7548, Ap
9	29	67.4	62	16 US-10-174-693-300	Sequence 300, App
10	29	67.4	223	15 US-10-156-761-8535	Sequence 8535, Ap
11	29	67.4	240	10 US-09-908-193-43	Sequence 43, Appl
12	29	67.4	240	11 US-09-823-187-74	Sequence 74, Appl
13	29	67.4	264	14 US-10-029-180-30	Sequence 30, Appl
14	29	67.4	417	11 US-09-972-288-25	Sequence 25, Appl
15	29	67.4	584	10 US-09-902-525-33	Sequence 33, Appl

16	29	67.4	944	14 US-10-029-180-26	Sequence 26, Appl
17	28	65.1	333	14 US-10-093-045-10	Sequence 10, Appl
18	28	65.1	333	15 US-10-093-246-10	Sequence 10, Appl
19	28	65.1	351	15 US-10-253-007-23	Sequence 23, Appl
20	28	65.1	353	15 US-10-253-007-32	Sequence 32, Appl
21	28	65.1	474	9 US-09-815-242-5389	Sequence 5389, Ap
22	28	65.1	474	9 US-09-815-242-12656	Sequence 12656, A
23	28	65.1	580	10 US-09-738-626-6264	Sequence 6264, Ap
24	28	65.1	677	15 US-10-156-761-10001	Sequence 10001, A
25	28	65.1	3798	14 US-10-014-717-6	Sequence 6, Appli
26	27	62.8	37	12 US-10-356-088-2	Sequence 2, Appli
27	27	62.8	37	12 US-10-356-088-25	Sequence 25, Appl
28	27	62.8	52	16 US-10-218-102-403	Sequence 403, App
29	27	62.8	57	11 US-09-764-891-4628	Sequence 4628, Ap
30	27	62.8	67	11 US-09-764-891-5011	Sequence 5011, Ap
31	27	62.8	67	15 US-10-205-428-423	Sequence 423, App
32	27	62.8	171	15 US-10-156-761-13832	Sequence 13832, A
33	27	62.8	179	16 US-10-174-693-299	Sequence 299, App
34	27	62.8	240	15 US-10-156-761-8139	Sequence 8139, Ap
35	27	62.8	247	10 US-09-908-193-42	Sequence 42, Appl
36	27	62.8	247	11 US-09-823-187-72	Sequence 72, Appl
37	27	62.8	251	9 US-09-835-788A-13	Sequence 13, Appl
38	27	62.8	251	12 US-10-175-042-13	Sequence 13, Appl
39	27	62.8	284	15 US-10-106-698-4666	Sequence 4666, Ap
40	27	62.8	305	15 US-10-156-761-11557	Sequence 11557, A
41	27	62.8	357	15 US-10-156-761-14681	Sequence 14681, A
42	27	62.8	383	15 US-10-224-446-2	Sequence 2, Appli
43	27	62.8	383	15 US-10-287-401-2	Sequence 2, Appli
44	27	62.8	390	12 US-10-032-585-7068	Sequence 7068, Ap
45	27	62.8	392	12 US-10-287-274-375	Sequence 375, App

## ALIGNMENTS

## RESULT 1

US-10-177-293-106

; Sequence 106, Application US/10177293

; Publication No. US20030124128A1

; GENERAL INFORMATION:

; APPLICANT: Lillie, James

; APPLICANT: Glatte, Karen

; APPLICANT: Zhao, Xumei

; APPLICANT: Gannavarpu, Manjula

; APPLICANT: Kamatkar, Shubhangi

; APPLICANT: Mertens, Maureen

; APPLICANT: Myer, Vic

; APPLICANT: Wang, Youzhen

; APPLICANT: Xu, Yongyao

; APPLICANT: Hoersch, Sebastian

; APPLICANT: Monahan, John

; APPLICANT: Meyers, Rachel E.

; APPLICANT: Bast Jr., Robert C.

; APPLICANT: Hortobagyi, Gabriel N.

; APPLICANT: Pusztai, Lajos

; APPLICANT: Meric, Funda

; APPLICANT: Sahin, Aysegul

; APPLICANT: Mills, Gordon B.

; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,

; FILE REFERENCE: MRI-038

; CURRENT FILING DATE: 2002-06-21

; PRIOR FILING DATE: 2002-06-21

; PRIOR FILING DATE: 2001-06-21

; PRIOR FILING DATE: 2001-06-21

; PRIOR FILING DATE: 2001-06-21

; PRIOR FILING DATE: 2001-06-21

; PRIOR FILING DATE: 2001-06-21

; PRIOR FILING DATE: 2001-06-21

; PRIOR FILING DATE: 2001-06-21

; PRIOR FILING DATE: 2001-06-21

; PRIOR FILING DATE: 2001-06-21

; PRIOR FILING DATE: 2001-06-21

; PRIOR FILING DATE: 2001-06-21

; PRIOR FILING DATE: 2001-06-21

; PRIOR FILING DATE: 2001-06-21

; PRIOR FILING DATE: 2001-06-21

; PRIOR FILING DATE: 2001-06-21

```

; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 106
; LENGTH: 1101
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 332
; OTHER INFORMATION: Xaa - Any Amino Acid
US-10-177-293-106

Query Match      74.4%; Score 32; DB 15; Length 1101;
Best Local Similarity 66.7%; Pred. No. 2.8e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 FLMAQGAML 9
Db      219 FLAQGASI 227

RESULT 2
US-09-925-301-1327
; Sequence 1327, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1327
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1327

Query Match      72.1%; Score 31; DB 9; Length 88;
Best Local Similarity 66.7%; Pred. No. 32;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 FLMAQGAML 9
Db      17 FLALGAMI 25

RESULT 3
US-10-106-698-4877
; Sequence 4877, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 4877

; LENGTH: 88
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-106-698-4877

Query Match      72.1%; Score 31; DB 15; Length 88;
Best Local Similarity 66.7%; Pred. No. 32;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 FLMAQGAML 9
Db      17 FLALGAMI 25

RESULT 4
US-10-156-761-14999
; Sequence 14999, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14999
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14999

Query Match      72.1%; Score 31; DB 15; Length 184;
Best Local Similarity 55.6%; Pred. No. 70;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      1 FLMAQGAML 9
Db      150 YLLAEGAMV 158

RESULT 5
US-10-181-590-8
; Sequence 8, Application US/10181590
; Publication No. US20030152949A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: BANDMAN, Olga
; APPLICANT: MATHUR, Preete
; APPLICANT: TANG, Y. Tom
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: YUE, Henry
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: WANG, Eureka
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: POLICKY, Jennifer L.
; TITLE OF INVENTION: PHOSPHATASES
; FILE REFERENCE: PI-0018 PCT
; CURRENT APPLICATION NUMBER: US/10/181,590
; CURRENT FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 60/177,719; 60/178,988; 60/184,959; 60/190,142
; PRIOR FILING DATE: 2000-01-21; 2000-01-28; 2000-02-25; 2000-03-17

```

```

; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PERL Program
; SEQ ID NO 8
; LENGTH: 402
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030152949A1 2522707CD1
US-10-181-590-8

Query Match          69.8%; Score 30; DB 12; Length 402;
Best Local Similarity 55.6%; Pred. No. 2.6e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 FLMAQGAML 9
||: ||:|
Db      351 FLLEQSGMI 359

RESULT 6
US-09-765-272-158
; Sequence 158, Application US/09765272
; Patent No. US20020061545A1
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,272
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 158:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 471 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 158:
US-09-765-272-158

Query Match          69.8%; Score 30; DB 9; Length 471;
Best Local Similarity 66.7%; Pred. No. 3e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 FLMAQGAML 9
||: ||:|
Db      151 FLNGQGAMM 159

; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PERL Program
; SEQ ID NO 8
; LENGTH: 402
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030152949A1 2522707CD1
US-10-181-590-8

Query Match          69.8%; Score 30; DB 12; Length 402;
Best Local Similarity 55.6%; Pred. No. 2.6e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 FLMAQGAML 9
||: ||:|
Db      351 FLLEQSGMI 359

RESULT 7
US-09-769-744A-98
; Sequence 98, Application US/09769744A
; Publication No. US20030134407A1
; GENERAL INFORMATION:
; APPLICANT: Le Page, Richard WF
; APPLICANT: Wells, Jeremy M
; APPLICANT: Hanniffy, Sean B
; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21122WO
; CURRENT APPLICATION NUMBER: US/09/769,744A
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: PCT/GB99/02452
; PRIOR FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: GB 9816336.3
; PRIOR FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: US 60/125329
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 98
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-769-744A-98

Query Match          69.8%; Score 30; DB 12; Length 490;
Best Local Similarity 66.7%; Pred. No. 3.1e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 FLMAQGAML 9
||: ||:|
Db      170 FLNGQGAMM 178

RESULT 8
US-10-032-585-7548
; Sequence 7548, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7548
; LENGTH: 1011
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-032-585-7548

Query Match          69.8%; Score 30; DB 12; Length 1011;
Best Local Similarity 62.5%; Pred. No. 6.7e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 FLMAQGAM 8
||: ||:|
Db      882 FLAEGSM 889

RESULT 9
US-10-174-693-300
; Sequence 300, Application US/10174693
; Publication No. US20030131373A1
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka

```

```

; TITLE OF INVENTION: Materials and Methods for the
; TITLE OF INVENTION: Modification of Plant Lignin Content
; FILE REFERENCE: 11000.1003c5
; CURRENT APPLICATION NUMBER: US/10/174,693
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US 08/975,316
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: US 08/713,000
; PRIOR FILING DATE: 1996-09-11
; PRIOR APPLICATION NUMBER: US 09/169,789
; PRIOR FILING DATE: 1998-10-09
; PRIOR APPLICATION NUMBER: US 09/615,192
; PRIOR FILING DATE: 2000-07-12
; NUMBER OF SEQ ID NOS: 407
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 300
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-10-174-693-300

Query Match 67.4%; Score 29; DB 16; Length 62;
Best Local Similarity 56.7%; Pred. No. 59;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLMAQGAML 9
Db 18 FCMAQGEVL 26

RESULT 10
US-10-156-761-8535
; Sequence 8535, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 8535
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-8535

Query Match 67.4%; Score 29; DB 15; Length 223;
Best Local Similarity 75.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LMAQGAML 9
Db 30 LVAQGTML 37

RESULT 11
US-09-908-193-43
; Sequence 43, Application US/09908193
; Publication No. US20020192748A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; APPLICANT: SHIMKETS, RICHARD A.

; TITLE OF INVENTION: Materials and Methods for the
; TITLE OF INVENTION: Modification of Plant Lignin Content
; FILE REFERENCE: 11000.1003c5
; CURRENT APPLICATION NUMBER: US/10/174,693
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US 08/975,316
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: US 08/713,000
; PRIOR FILING DATE: 1996-09-11
; PRIOR APPLICATION NUMBER: US 09/169,789
; PRIOR FILING DATE: 1998-10-09
; PRIOR APPLICATION NUMBER: US 09/615,192
; PRIOR FILING DATE: 2000-07-12
; NUMBER OF SEQ ID NOS: 407
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 300
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-10-174-693-300

Query Match 67.4%; Score 29; DB 16; Length 62;
Best Local Similarity 56.7%; Pred. No. 59;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLMAQGAML 9
Db 18 FCMAQGEVL 26

RESULT 10
US-10-156-761-8535
; Sequence 8535, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 8535
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-8535

Query Match 67.4%; Score 29; DB 15; Length 223;
Best Local Similarity 75.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LMAQGAML 9
Db 30 LVAQGTML 37

RESULT 11
US-09-908-193-43
; Sequence 43, Application US/09908193
; Publication No. US20020192748A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; APPLICANT: SHIMKETS, RICHARD A.

; APPLICANT: ZERHUSEN, BRYAN
; APPLICANT: MALYANKAR, URIEL M.
; APPLICANT: PADIGARU, MURALIDHARA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES ENCODED THEREBY
; FILE REFERENCE: 21402-062
; CURRENT APPLICATION NUMBER: US/09/908,193
; CURRENT FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: 60/220,273
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: 60/221,650
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/221,233
; PRIOR FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: 60/220,912
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/218,875
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/218,870
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/218,901
; PRIOR FILING DATE: 2000-07-18
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-908-193-43

Query Match 67.4%; Score 29; DB 10; Length 240;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLMAQGAML 9
Db 59 FLIAGAVL 67

RESULT 12
US-09-823-187-74
; Sequence 74, Application US/09823187
; Publication No. US20030096952A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Catherine
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Liu, Xiaohong
; APPLICANT: Majumder, Kumud
; APPLICANT: Padigaru, Muralidhar
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard A
; APPLICANT: Spaderna, Steven K
; APPLICANT: Spytek, Kimberly
; APPLICANT: Taupier, Raymond J
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-745
; CURRENT APPLICATION NUMBER: US/09/823,187
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/193,339
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/193,205
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/195,343
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/195,088
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,005
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,792
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: 60/196,556
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: 60/197,081
; PRIOR FILING DATE: 2000-04-13
```

; PRIOR APPLICATION NUMBER: 60/197,525  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/197,087  
; PRIOR FILING DATE: 2000-04-14  
; NUMBER OF SEQ ID NOS: 103  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 74  
; LENGTH: 240  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-823-187-74

Query Match 67.4%; Score 29; DB 11; Length 240;  
Best Local Similarity 66.7%; Pred. No. 2.4e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9  
||| |||  
Db 59 FLIAGAVL 67

## RESULT 13

US-10-029-180-30  
; Sequence 30, Application US/10029180  
; Publication No. US20020182708A1  
; GENERAL INFORMATION:  
; APPLICANT: Cali, Brian M.  
; APPLICANT: Holtzman, Doug  
; APPLICANT: Madden, Kevin T.  
; APPLICANT: Milna, G. Todd  
; APPLICANT: Sherman, Amir  
; APPLICANT: Silva, Jeffrey C.  
; APPLICANT: Trueheart, Josh  
; APPLICANT: Zhang, Lixin  
; TITLE OF INVENTION: No. US20020182708A1el Regulators of Fungal Gene Expression  
; FILE REFERENCE: MIC-004  
; CURRENT APPLICATION NUMBER: US/10/029,180  
; CURRENT FILING DATE: 2001-12-22  
; PRIOR APPLICATION NUMBER: US 60/257,431  
; PRIOR FILING DATE: 2000-12-22  
; NUMBER OF SEQ ID NOS: 138  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 30  
; LENGTH: 264  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: fungal gene  
US-10-029-180-30

Query Match 67.4%; Score 29; DB 14; Length 264;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 MAQAM 8  
|||||  
Db 164 MAQAM 169

## RESULT 14

US-09-972-268-25  
; Sequence 25, Application US/09972268  
; Publication No. US20030044893A1  
; GENERAL INFORMATION:  
; APPLICANT: Baum, Peter R.  
; APPLICANT: Fanslow, William C.  
; APPLICANT: Lofton, Timothy E.  
; APPLICANT: Sorensen, Eric A.  
; APPLICANT: Youakim, Adel  
; TITLE OF INVENTION: NECTIN POLYPEPTIDES, POLYNUCLEOTIDES, METHODS OF MAKING AND USE  
; FILE REFERENCE: 3101-A  
; CURRENT APPLICATION NUMBER: US/09/972,268  
; CURRENT FILING DATE: 2001-10-05

; PRIOR APPLICATION NUMBER: 60/238,557  
; PRIOR FILING DATE: 2000-10-05  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 25  
; LENGTH: 417  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-09-972-268-25

Query Match 67.4%; Score 29; DB 11; Length 417;  
Best Local Similarity 66.7%; Pred. No. 4.3e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9  
||| |||  
Db 289 FAVAQAGL 297

## RESULT 15

US-09-902-525-33  
; Sequence 33, Application US/09902525  
; Patent No. US20020164751A1  
; GENERAL INFORMATION:  
; APPLICANT: Recombinant Biocatalysis, Inc.  
; TITLE OF INVENTION: THERMOSTABLE PHOSPHATASES  
; NUMBER OF SEQUENCES: 54  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows 95  
; SOFTWARE: FastSeq for Windows Version 2.0b  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/902,525  
; FILING DATE: 09-Jul-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/202,681A  
; FILING DATE: 23-Dec-1999  
; APPLICATION NUMBER: 08/666,857  
; FILING DATE: 19-JUN-1996  
; APPLICATION NUMBER: 60/033,752  
; FILING DATE: 19-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haile, Ph.D., Lisa A.  
; REGISTRATION NUMBER: 38,347  
; REFERENCE/DOCKET NUMBER: 09010/015W01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619/678-5070  
; TELEFAX: 619/678-5099  
; TELEX: <Unknown>  
; INFORMATION FOR SEQ ID NO: 33:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 584 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
; SEQUENCE DESCRIPTION: SEQ ID NO: 33:  
US-09-902-525-33

Query Match 67.4%; Score 29; DB 10; Length 584;  
Best Local Similarity 55.6%; Pred. No. 6.1e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9

us-09-807-512-24.rapb

Tue Oct 7 15:52:36 2003

Db           |||:|:|:  
          381 FLWEGGMV 389

Search completed: October 7, 2003, 13:58:56  
Job time : 10.7975 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:25:43 ; Search time 70.8608 Seconds  
(without alignments)  
115.568 Million cell updates/sec

Title: US-09-807-512-24  
Perfect score: 43  
Sequence: 1 FLMAQGAML 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues

Total number of hits satisfying chosen parameters: 5728757

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main:\*  
1: /cgn2\_6/ptodata/1/paa/PCTUS\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep.\*  
8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep.\*  
9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep.\*  
10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep.\*  
11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep.\*  
12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep.\*  
13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep.\*  
14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep.\*  
15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep.\*  
16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep.\*  
17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep.\*  
18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep.\*  
19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep.\*  
20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep.\*  
21: /cgn2\_6/ptodata/1/paa/US097A\_COMB.pep.\*  
22: /cgn2\_6/ptodata/1/paa/US097B\_COMB.pep.\*  
23: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep.\*  
24: /cgn2\_6/ptodata/1/paa/US099A\_COMB.pep.\*  
25: /cgn2\_6/ptodata/1/paa/US099B\_COMB.pep.\*  
26: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep.\*  
27: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep.\*  
28: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep.\*  
29: /cgn2\_6/ptodata/1/paa/US103\_COMB.pep.\*  
30: /cgn2\_6/ptodata/1/paa/US104\_COMB.pep.\*  
31: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep.\*  
32: /cgn2\_6/ptodata/1/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	100.0	9	23	US-09-807-512-24
					Sequence 24, Appl

2	43	100.0	25	19	US-09-529-206-56	Sequence 56, Appl
3	43	100.0	25	19	US-09-529-206A-56	Sequence 56, Appl
4	43	100.0	25	19	US-09-529-206B-56	Sequence 56, Appl
5	43	100.0	25	19	US-09-529-206D-56	Sequence 56, Appl
6	43	100.0	58	19	US-09-529-206-5	Sequence 5, Appl
7	43	100.0	58	19	US-09-529-206A-5	Sequence 5, Appl
8	43	100.0	58	19	US-09-529-206B-5	Sequence 5, Appl
9	43	100.0	58	19	US-09-529-206D-5	Sequence 5, Appl
10	43	100.0	58	23	US-09-807-512-10	Sequence 2, Appl
11	43	100.0	109	23	US-09-807-512-2	Sequence 2, Appl
12	33	76.7	256	21	US-09-733-089-16050	Sequence 16050, A
13	33	76.7	256	23	US-09-816-660-16050	Sequence 16050, A
14	33	76.7	297	18	US-09-489-039A-10089	Sequence 10089, A
15	33	76.7	297	30	US-10-446-203-10089	Sequence 10089, A
16	32	74.4	181	1	PCT-US01-08631-32398	Sequence 32398, A
17	32	74.4	350	22	US-09-791-537-8615	Sequence 8615, Ap
18	32	74.4	350	22	US-09-791-537-37430	Sequence 17430, A
19	32	74.4	406	24	US-09-937-059-23	Sequence 23, Appl
20	32	74.4	406	26	US-10-094-749-2164	Sequence 2164, Ap
21	32	74.4	406	32	US-60-139-565-10	Sequence 10, Appl
22	32	74.4	1101	1	PCT-US02-19669A-106	Sequence 106, App
23	32	74.4	1101	1	PCT-US02-19669A-106	Sequence 106, App
24	32	74.4	1101	27	US-10-177-293-106	Sequence 106, App
25	32	74.4	1101	30	US-10-408-765A-2181	Sequence 2181, Ap
26	32	74.4	1101	32	US-60-389-987-2181	Sequence 2181, Ap
27	32	74.4	1101	32	US-60-412-418-2181	Sequence 2181, Ap
28	32	74.4	1360	1	PCT-US02-09288-12	Sequence 12, Appl
29	32	74.4	1471	1	PCT-US01-14827-11753	Sequence 11753, A
30	31	72.1	88	1	PCT-US00-05882-1327	Sequence 1327, Ap
31	31	72.1	88	1	PCT-US00-26524B-4867	Sequence 4867, Ap
32	31	72.1	88	24	US-09-925-301-1327	Sequence 1327, Ap
33	31	72.1	88	27	US-10-106-698-4877	Sequence 4877, Ap
34	31	72.1	184	27	US-10-156-761-14999	Sequence 14999, A
35	31	72.1	221	32	US-60-173-464-23537	Sequence 23537, A
36	31	72.1	222	20	US-09-614-150-30072	Sequence 30072, A
37	31	72.1	222	32	US-60-191-637-29620	Sequence 29620, A
38	31	72.1	222	32	US-60-191-681-23935	Sequence 23935, A
39	31	72.1	227	30	US-10-437-963-129549	Sequence 129549, A
40	31	72.1	423	19	US-09-513-996A-36620	Sequence 36620, A
41	31	72.1	423	19	US-09-513-996A-36620	Sequence 36620, A
42	31	72.1	423	21	US-09-708-427-25083	Sequence 25083, A
43	31	72.1	423	21	US-09-708-427-25087	Sequence 25087, A
44	31	72.1	430	18	US-09-417-507-36755	Sequence 36755, A
45	31	72.1	439	19	US-09-513-996A-36619	Sequence 36619, A

ALIGNMENTS

RESULT 1  
US-09-807-512-24  
; Sequence 24, Application US/09807512  
; GENERAL INFORMATION:  
; APPLICANT: Schrier, Peter I.  
; APPLICANT: Aarnoudse, Corlien  
; APPLICANT: Heider, Karl-Heinz  
; APPLICANT: Klade, Christoph  
; TITLE OF INVENTION: Camel, An Alternative Translation Product of the Tumor  
; TITLE OF INVENTION: Antigen-Lage 1  
; FILE REFERENCE: 0652.2200000  
; CURRENT APPLICATION NUMBER: US/09/807,512  
; CURRENT FILING DATE: 2001-04-16  
; PRIOR APPLICATION NUMBER: PCT/EP99/07832  
; PRIOR FILING DATE: 1999-10-15  
; PRIOR APPLICATION NUMBER: EP 98119583.7  
; PRIOR FILING DATE: 1998-10-16  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 24  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-807-512-24

```
Query Match      100.0%; Score 43; DB 23; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.3e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLMAQGAML 9
   |||||
Db 1 FLMAQGAML 9

RESULT 2
US-09-529-206-56
; Sequence 56, Application US/09529206
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; FILE REFERENCE: 20264269US1
; CURRENT APPLICATION NUMBER: US/09/529,206B
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: PCT/US98/19609
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206-56

Query Match      100.0%; Score 43; DB 19; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLMAQGAML 9
   |||||
Db 10 FLMAQGAML 18

RESULT 3
US-09-529-206A-56
; Sequence 56, Application US/09529206A
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; FILE REFERENCE: 20264269US1
; CURRENT APPLICATION NUMBER: US/09/529,206A
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: PCT/US98/19609
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206A-56

Query Match      100.0%; Score 43; DB 19; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLMAQGAML 9
   |||||
Db 10 FLMAQGAML 18

RESULT 4
US-09-529-206B-56
; Sequence 56, Application US/09529206B
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; FILE REFERENCE: 20264269US1
; CURRENT APPLICATION NUMBER: US/09/529,206B
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: PCT/US98/19609
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206B-56

Query Match      100.0%; Score 43; DB 19; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLMAQGAML 9
   |||||
Db 10 FLMAQGAML 18

RESULT 5
US-09-529-206D-56
; Sequence 56, Application US/09529206D
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; FILE REFERENCE: 218791
; CURRENT APPLICATION NUMBER: US/09/529,206D
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: PCT/US98/19609
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206D-56

Query Match      100.0%; Score 43; DB 19; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLMAQGAML 9
   |||||
Db 10 FLMAQGAML 18

RESULT 6
US-09-529-206-5
; Sequence 5, Application US/09529206
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America as
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; FILE REFERENCE: 20264269PC
; CURRENT APPLICATION NUMBER: US/09/529,206
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US60/061,428
```



; PRIOR FILING DATE: 1997-10-08  
; NUMBER OF SEQ ID NOS: 106  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 58  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-206-5

Query Match 100.0%; Score 43; DB 19; Length 58;  
Best Local Similarity 100.0%; Pred. No. 0.42;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9  
Db 10 FLMAQGAML 18

RESULT 7  
US-09-529-206A-5  
; Sequence 5, Application US/09529206A  
; GENERAL INFORMATION:  
; APPLICANT: Wang Rong, Fu  
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene  
; FILE REFERENCE: 20264269US1  
; CURRENT FILING DATE: 2000-06-13  
; PRIOR APPLICATION NUMBER: US/09/529,206A  
; PRIOR FILING DATE: 1998-09-21  
; PRIOR APPLICATION NUMBER: PCT/US98/19609  
; PRIOR FILING DATE: 1998-09-21  
; PRIOR APPLICATION NUMBER: US60/061,428  
; PRIOR FILING DATE: 1997-10-08  
; NUMBER OF SEQ ID NOS: 127  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 58  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-206A-5

Query Match 100.0%; Score 43; DB 19; Length 58;  
Best Local Similarity 100.0%; Pred. No. 0.42;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9  
Db 10 FLMAQGAML 18

RESULT 8  
US-09-529-206B-5  
; Sequence 5, Application US/09529206B  
; GENERAL INFORMATION:  
; APPLICANT: Wang Rong, Fu  
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene  
; FILE REFERENCE: 20264269US1  
; CURRENT FILING DATE: 2000-06-13  
; PRIOR APPLICATION NUMBER: PCT/US98/19609  
; PRIOR FILING DATE: 1998-09-21  
; PRIOR APPLICATION NUMBER: US60/061,428  
; PRIOR FILING DATE: 1997-10-08  
; NUMBER OF SEQ ID NOS: 127  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 58  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-206B-5

Query Match 100.0%; Score 43; DB 19; Length 58;  
Best Local Similarity 100.0%; Pred. No. 0.42;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9  
Db 10 FLMAQGAML 18

RESULT 9  
US-09-529-206D-5  
; Sequence 5, Application US/09529206D  
; GENERAL INFORMATION:  
; APPLICANT: Wang Rong, Fu  
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene  
; FILE REFERENCE: 218791  
; CURRENT FILING DATE: 2000-06-13  
; PRIOR APPLICATION NUMBER: PCT/US98/19609  
; PRIOR FILING DATE: 1998-09-21  
; PRIOR APPLICATION NUMBER: US60/061,428  
; PRIOR FILING DATE: 1997-10-08  
; NUMBER OF SEQ ID NOS: 127  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 58  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-206D-5

Query Match 100.0%; Score 43; DB 19; Length 58;  
Best Local Similarity 100.0%; Pred. No. 0.42;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9  
Db 10 FLMAQGAML 18

RESULT 10  
US-09-807-512-10  
; Sequence 10, Application US/09807512  
; GENERAL INFORMATION:  
; APPLICANT: Schrier, Peter I.  
; APPLICANT: Aarnoudse, Corlien  
; APPLICANT: Heider, Karl-Heinz  
; APPLICANT: Klade, Christoph  
; TITLE OF INVENTION: Camel, An Alternative Translation Product of the Tumor  
; FILE REFERENCE: 0652.2200000  
; CURRENT FILING DATE: 2001-04-16  
; PRIOR APPLICATION NUMBER: PCT/EP99/07832  
; PRIOR FILING DATE: 1999-10-15  
; PRIOR APPLICATION NUMBER: EP 98119583.7  
; PRIOR FILING DATE: 1998-10-16  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 58  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-807-512-10

Query Match 100.0%; Score 43; DB 23; Length 58;  
Best Local Similarity 100.0%; Pred. No. 0.42;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9  
Db 10 FLMAQGAML 18

```

RESULT 11
US-09-807-512-2
; Sequence 2, Application US/09807512
; GENERAL INFORMATION:
; APPLICANT: Schrier, Peter I.
; APPLICANT: Aarnoudse, Corlien
; APPLICANT: Heider, Karl-Heinz
; APPLICANT: Klade, Christoph
; TITLE OF INVENTION: Camel, An Alternative Translation Product of the Tumor
; TITLE OF INVENTION: Antigen-Lage 1
; FILE REFERENCE: 0652.2200000
; CURRENT APPLICATION NUMBER: US/09/807,512
; CURRENT FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: PCT/EP99/07832
; PRIOR FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: EP 98119583.7
; PRIOR FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-807-512-2

Query Match 100.0%; Score 43; DB 23; Length 109;
Best Local Similarity 100.0%; Pred. No. 0.82;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9
Db 10 FLMAQGAML 18

RESULT 12
US-09-733-089-16050
; Sequence 16050, Application US/09733089
; GENERAL INFORMATION:
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Lutfiyva, Linda L.
; APPLICANT: McIninch, James
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; TITLE OF INVENTION: Transcription In Plants
; FILE REFERENCE: 38-21(15300)D
; CURRENT APPLICATION NUMBER: US/09/733,089
; CURRENT FILING DATE: 2000-12-11
; PRIOR APPLICATION NUMBER: US/09/474,435
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: US/09/654,617
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US/09/620,392
; PRIOR FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 24143
; SEQ ID NO 16050
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-733-089-16050

Query Match 76.7%; Score 33; DB 21; Length 256;
Best Local Similarity 75.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAQGAM 8
Db 84 FLMAQSGM 91

RESULT 13
US-09-816-660-16050
; Sequence 16050, Application US/09816660
; GENERAL INFORMATION:
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Lutfiyva, Linda L.
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; TITLE OF INVENTION: Transcription In Plants
; FILE REFERENCE: 38-21(15300)D
; CURRENT APPLICATION NUMBER: US/09/816,660
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US/09/474,435
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: US/09/654,617
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US/09/733,089
; PRIOR FILING DATE: 2000-12-11
; PRIOR APPLICATION NUMBER: )
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US/09/620,392
; PRIOR FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 24143
; SEQ ID NO 16050
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-816-660-16050

Query Match 76.7%; Score 33; DB 23; Length 256;
Best Local Similarity 75.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAQGAM 8
Db 84 FLMAQSGM 91

RESULT 14
US-09-489-039A-10089
; Sequence 10089, Application US/09489039A
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709,2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10089
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10089

Query Match 76.7%; Score 33; DB 18; Length 297;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LMAQGAM 8
Db 19 LMAQGAM 25

RESULT 15
US-10-446-203-10089
; Sequence 10089, Application US/10446203
; GENERAL INFORMATION:
```

; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/10/446,203  
; CURRENT FILING DATE: 2003-05-27  
; PRIOR APPLICATION NUMBER: US/09/489,039  
; PRIOR FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 10089  
; LENGTH: 297  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-10-446-203-10089

Query Match 76.7%; Score 33; DB 30; Length 297;  
Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LMAQGAM 8  
| | | | | | | |  
DB 19 LMAQGAM 25

Search completed: October 7, 2003, 13:55:00  
Job time : 71.8608 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:26:33 ; Search time 2.8481 Seconds  
(without alignments)  
92.039 Million cell updates/sec

Title: US-09-807-512-24  
Perfect score: 43  
Sequence: 1 FLMAQGAML 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 136899 seqs, 29126274 residues

Total number of hits satisfying chosen parameters: 136899

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*  
1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	43	100.0	30	US-10-296-734-1422	Sequence 1422, Ap
2	43	100.0	58	US-10-296-734-833	Sequence 833, App
3	43	100.0	3541	US-10-296-734-1454	Sequence 1454, Ap
4	30	69.8	434	US-10-296-715-742	Sequence 742, App
5	29	67.4	131	US-10-425-114A-42299	Sequence 42299, A
6	29	67.4	364	PCT-US03-28227-5309	Sequence 5309, Ap
7	29	67.4	372	PCT-US03-28227-4488	Sequence 4488, Ap
8	28	65.1	166	US-09-897-516A-8256	Sequence 8256, Ap
9	27	62.8	95	US-10-425-114A-57389	Sequence 57389, A
10	27	62.8	186	US-10-425-114A-56900	Sequence 56900, A
11	27	62.8	200	US-09-674-546A-3007	Sequence 3007, Ap
12	27	62.8	207	US-09-674-546A-3009	Sequence 3009, Ap
13	27	62.8	207	US-09-674-546A-3011	Sequence 3011, Ap
14	27	62.8	212	PCT-US03-10753-181	Sequence 181, App
15	27	62.8	212	US-10-410-031-181	Sequence 181, App
16	27	62.8	366	US-10-425-114A-52567	Sequence 52567, A
17	27	62.8	433	US-10-425-114A-41940	Sequence 41940, A
18	27	62.8	445	PCT-US30-21378-35	Sequence 35, Appl
19	27	62.8	445	US-09-897-516A-7647	Sequence 7647, Ap
20	27	62.8	447	PCT-US30-21378-36	Sequence 36, Appl
21	27	62.8	517	US-60-500-337-1727	Sequence 1727, Ap
22	27	62.8	521	US-60-500-337-1726	Sequence 1726, Ap
23	27	62.8	549	US-10-425-114A-63960	Sequence 63960, A
24	27	62.8	928	PCT-US03-17937-30	Sequence 30, Appl
25	27	62.8	928	US-60-485-450-1542	Sequence 1542, Ap
26	27	62.8	1856	PCT-US03-27401-244	Sequence 244, App

27	26	60.5	37	5	US-09-653-465B-6	Sequence 6, Appli
28	26	60.5	110	6	US-10-654-578-27	Sequence 27, Appl
29	26	60.5	115	1	PCT-US02-41612A-578	Sequence 578, App
30	26	60.5	141	6	US-10-425-114A-49334	Sequence 49334, A
31	26	60.5	201	7	US-60-487-610-2212	Sequence 2212, Ap
32	26	60.5	201	7	US-60-485-450-1410	Sequence 1410, Ap
33	26	60.5	218	5	US-09-840-243C-19	Sequence 19, Appl
34	26	60.5	220	5	US-09-840-243C-18	Sequence 18, Appl
35	26	60.5	237	7	US-60-500-315-1295	Sequence 1295, Ap
36	26	60.5	241	7	US-60-495-114-1881	Sequence 1881, Ap
37	26	60.5	313	7	US-60-497-790-82	Sequence 82, Appl
38	26	60.5	314	1	PCT-US03-18840-176	Sequence 176, App
39	26	60.5	318	6	US-10-343-650A-124	Sequence 124, App
40	26	60.5	365	6	US-10-343-650A-122	Sequence 122, App
41	26	60.5	430	7	US-60-487-610-2213	Sequence 2213, Ap
42	26	60.5	430	7	US-60-485-450-1411	Sequence 1411, Ap
43	26	60.5	430	7	US-60-495-114-1882	Sequence 1882, Ap
44	26	60.5	437	6	US-10-425-114A-46385	Sequence 46385, A
45	26	60.5	438	6	US-10-425-114A-40595	Sequence 40595, A

ALIGNMENTS

RESULT 1  
US-10-296-734-1422  
; Sequence 1422, Application US/10296734  
; GENERAL INFORMATION:  
; APPLICANT: Thompson, Scott A  
; TITLE OF INVENTION: Synthetic molecules and uses therefor  
; FILE REFERENCE: Savine  
; CURRENT APPLICATION NUMBER: US/10/296,734  
; CURRENT FILING DATE: 2003-08-04  
; PRIOR APPLICATION NUMBER: AU PQ7761/00  
; PRIOR FILING DATE: 2000-05-26  
; NUMBER OF SEQ ID NOS: 1507  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1422  
; LENGTH: 30  
; TYPE: PRT  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: NYSolb segment 1  
US-10-296-734-1422

Query Match 100.0%; Score 43; DB 6; Length 30;  
Best Local Similarity 100.0%; Pred. No. 0.0065;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLMAQGAML 9  
Db 12 FLMAQGAML 20

RESULT 2  
US-10-296-734-833  
; Sequence 833, Application US/10296734  
; GENERAL INFORMATION:  
; APPLICANT: Thompson, Scott A  
; TITLE OF INVENTION: Synthetic molecules and uses therefor  
; FILE REFERENCE: Savine  
; CURRENT APPLICATION NUMBER: US/10/296,734  
; CURRENT FILING DATE: 2003-08-04  
; PRIOR APPLICATION NUMBER: AU PQ7761/00  
; PRIOR FILING DATE: 2000-05-26  
; NUMBER OF SEQ ID NOS: 1507  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 833  
; LENGTH: 58  
; TYPE: PRT  
; ORGANISM: Artificial

Tue Oct 7 15:52:37 2003

us-09-807-512-24.rapn

```

; FEATURE:
; OTHER INFORMATION: NYN501b consensus polypeptide
US-10-296-734-833

Query Match      100.0%; Score 43; DB 6; Length 58;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FLMAQGAML 9
Db      10 FLMAQGAML 18

RESULT 3
US-10-296-734-1454
; Sequence 1454, Application US/10296734
; GENERAL INFORMATION:
; APPLICANT: Thompson, Scott A
; APPLICANT: Ramshaw, Ian A
; TITLE OF INVENTION: Synthetic molecules and uses therefor
; FILE REFERENCE: Savine
; CURRENT APPLICATION NUMBER: US/10/296.734
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: AU PQ7761/00
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1454
; LENGTH: 3541
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Melanoma cancer specific savine
US-10-296-734-1454

Query Match      100.0%; Score 43; DB 6; Length 3541;
Best Local Similarity 100.0%; Pred. No. 0.63;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FLMAQGAML 9
Db      1763 FLMAQGAML 1771

RESULT 4
US-10-296-115-742
; Sequence 742, Application US/10296115
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: US/10/296.115
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/488.725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552.317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 742
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-296-115-742

Query Match      69.8%; Score 30; DB 6; Length 434;
Best Local Similarity 55.6%; Pred. No. 46;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 FLMAQGAML 9
Db      383 FLLEQSGMI 391

; FEATURE:
; OTHER INFORMATION: NYN501b consensus polypeptide
US-10-425-114A-42299
; Sequence 42299, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425.114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 42299
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3356-044-H6_FLI.pep
US-10-425-114A-42299

Query Match      67.4%; Score 29; DB 6; Length 131;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 MAQGAM 8
Db      85 MAQGAM 90

RESULT 5
PCT-US03-28227-5309
; Sequence 5309, Application PC/TUS0328227
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; SCHMIDT, Jeanette P.;
; APPLICANT: WRIGHT, Rachel J.; BRUNS, Christopher M.;
; APPLICANT: MARJANOVIC, Mirjana M.; SHEN, Fan;
; APPLICANT: HARTSHORNE, Toinette A.; SUCHOROLSKI, Martin;
; APPLICANT: ALTUS, Christina M.; PITTS, Steven J.;
; APPLICANT: ELDER, Linda V.; MOONEY, Elizabeth M.;
; APPLICANT: DELEGEANE, Angelo M.; REDDY, Thirupathi P.;
; APPLICANT: BANVILLE, Steven C.; BLANCHARD, John L.;
; APPLICANT: STEVENS, Kristian A.; WANG, Xinhao;
; APPLICANT: PANZER, Scott R.; GERSTIN, Edward H., Jr.;
; APPLICANT: AU, Alan P.; CAREYNA H.; ANDERSON, Scott E.;
; APPLICANT: PERALTA, Careyna H.; SHEN, Edward J.;
; APPLICANT: RIOUX, Pierre; STUVE, Laura L.;
; APPLICANT: WU, Mingham C.; SPIRO, Peter A.;
; APPLICANT: LAGACE, Robert E.; WINGROVE, James A.;
; APPLICANT: STEWART, Elizabeth A.; KIRTON, Edward;
; APPLICANT: VITT, Ursula A.; KWONG, Mary;
; APPLICANT: XU, Yuming; Jennifer L.; HURWITZ, Bonnie L.;
; APPLICANT: POLICKY, Jennifer L.; JACKSON, Jennifer L.;
; APPLICANT: MA, Yan; JACKSON, Jennifer L.;
; APPLICANT: GIETZEN, Darryl; PATURY, Srikanth;
; APPLICANT: SHI, Xiaobing; SUAREZ, Charlyn J.
; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PN-0100 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/28227
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: US 60/410,260
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/410,259
; PRIOR FILING DATE: 2002-09-12
; NUMBER OF SEQ ID NOS: 5444
; SOFTWARE: PERL Program
; SEQ ID NO 5309
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo sapiens

```

;  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No: 960834.Pt34p  
PCT-US03-28227-5309

Query Match 67.4%; Score 29; DB 1; Length 364;  
Best Local Similarity 66.7%; Pred. No. 64;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9  
| : |||||  
Db 289 FAVAQAQL 297

## RESULT 7

PCT-US03-28227-4488  
; Sequence 4488, Application PC/TUS0328227  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE CORPORATION; SCHMIDT, Jeanette P.;  
; APPLICANT: WRIGHT, Rachel J.; BRUNS, Christopher M.;  
; APPLICANT: MARJANOVIC, Mirjana M.; SHEN, Fan;  
; APPLICANT: HARTSHORNE, Toinette A.; SUCHOROLSKI, Martin;  
; APPLICANT: ALTUS, Christina M.; PITTS, Steven J.;  
; APPLICANT: ELDER, Linda V.; MOONEY, Elizabeth M.;  
; APPLICANT: DELEGEANE, Angelo M.; PANESAR, Iqbal S.;  
; APPLICANT: BANYILLE, Steven C.; REDDY, Thirupathi P.;  
; APPLICANT: STEVENS, Kristian A.; BLANCHARD, John L.;  
; APPLICANT: PANZER, Scott R.; WANG, Xinhao;  
; APPLICANT: AU, Alan P.; GERSTIN, Edward H., Jr.;  
; APPLICANT: PERALTA, Careyana H.; ANDERSON, Scott E.;  
; APPLICANT: RIOUX, Pierre; SHEN, Edward J.;  
; APPLICANT: WU, Mingham C.; STUVE, Laura L.;  
; APPLICANT: LAGACE, Robert E.; SPIRO, Peter A.;  
; APPLICANT: STEWART, Elizabeth A.; WINGROVE, James A.;  
; APPLICANT: VITT, Ursula A.; KIRTON, Edward;  
; APPLICANT: XU, Yuming; KWONG, Mary;  
; APPLICANT: POLICKY, Jennifer L.; HURWITZ, Bonnie L.;  
; APPLICANT: MA, Yan; JACKSON, Jennifer L.;  
; APPLICANT: GIETZEN, Darryl; PATURY, Srikanth;  
; APPLICANT: SHI, Xiaobing; SUAREZ, Charlyn J.  
; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: PN-0100 PCT  
; CURRENT APPLICATION NUMBER: PCT/US03/28227  
; CURRENT FILING DATE: 2003-09-12  
; PRIOR APPLICATION NUMBER: US 60/410,260  
; PRIOR FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: US 60/410,259  
; PRIOR FILING DATE: 2002-09-12  
; NUMBER OF SEQ ID NOS: 544  
; SOFTWARE: PERL Program  
; SEQ ID NO 4488  
; LENGTH: 372  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No: 960834.Pt33p  
PCT-US03-28227-4488

Query Match 67.4%; Score 29; DB 1; Length 372;  
Best Local Similarity 66.7%; Pred. No. 65;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9  
| : |||||  
Db 289 FAVAQAQL 297

## RESULT 8

US-09-897-516A-8256  
; Sequence 8256, Application US/09897516A  
; GENERAL INFORMATION:  
; APPLICANT: Corbin, David R.

; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Huesing, Joseph E.  
; APPLICANT: Malvar, Thomas M.  
; APPLICANT: Krasomil-Osterfeld, Karina C.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Spiridonov, Sergei  
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof  
; FILE REFERENCE: 38-21(51847)B  
; CURRENT APPLICATION NUMBER: US/09/897,516A  
; CURRENT FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/215,161  
; PRIOR FILING DATE: 2000-06-30  
; NUMBER OF SEQ ID NOS: 8415  
; SEQ ID NO 8256  
; LENGTH: 166  
; TYPE: PRT  
; ORGANISM: Xenorhabdus sp.  
US-09-897-516A-8256

Query Match 65.1%; Score 28; DB 5; Length 166;  
Best Local Similarity 55.6%; Pred. No. 49;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLMAQGAML 9  
||| : |||  
Db 22 FLMKEGQMM 30

## RESULT 9

US-10-425-114A-57389  
; Sequence 57389, Application US/10425114A  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114A  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 57389  
; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-2MFLMO17308D02\_FLI.pep  
US-10-425-114A-57389

Query Match 62.8%; Score 27; DB 6; Length 95;  
Best Local Similarity 62.5%; Pred. No. 47;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 LMAQGAML 9  
|| : |||  
Db 4 LMLQGLI 11

## RESULT 10

US-10-425-114A-56900  
; Sequence 56900, Application US/10425114A  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

Tue Oct 7 15:52:37 2003

us-09-807-512-24.rapn

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53313)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,114A  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 73128  
 ; SEQ ID NO 56900  
 ; LENGTH: 186  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: UC-2MELMO17048F09\_FLI.pep  
 ; US-10-425-114A-56900

Query Match 62.8%; Score 27; DB 6; Length 186;  
 Best Local Similarity 71.4%; Pred. No. 89;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LMAQAM 8  
 Db 22 LLAQAL 28

RESULT 11  
 US-09-674-546A-3007  
 ; Sequence 3007, Application US/09674546A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Institute for Genomic Research  
 ; APPLICANT: Chiron Corporation  
 ; TITLE OF INVENTION: Neisseria Meningitidis Antigens and Compositions  
 ; FILE REFERENCE: CHIR-0334  
 ; CURRENT APPLICATION NUMBER: US/09/674,546A  
 ; CURRENT FILING DATE: 1999-04-30  
 ; NUMBER OF SEQ ID NOS: 3264  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 3007  
 ; LENGTH: 200  
 ; TYPE: PRT  
 ; ORGANISM: Neisseria gonorrhoeae  
 ; US-09-674-546A-3007

Query Match 62.8%; Score 27; DB 5; Length 200;  
 Best Local Similarity 55.6%; Pred. No. 95;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLMAQAML 9  
 Db 6 FLGAGALL 14

RESULT 12  
 US-09-674-546A-3009  
 ; Sequence 3009, Application US/09674546A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Institute for Genomic Research  
 ; APPLICANT: Chiron Corporation  
 ; TITLE OF INVENTION: Neisseria Meningitidis Antigens and Compositions  
 ; FILE REFERENCE: CHIR-0334  
 ; CURRENT APPLICATION NUMBER: US/09/674,546A  
 ; CURRENT FILING DATE: 1999-04-30  
 ; NUMBER OF SEQ ID NOS: 3264  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 3009  
 ; LENGTH: 207  
 ; TYPE: PRT  
 ; ORGANISM: Neisseria meningitidis  
 ; US-09-674-546A-3009

Query Match 62.8%; Score 27; DB 5; Length 207;  
 Best Local Similarity 55.6%; Pred. No. 98;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLMAQAML 9  
 Db 6 FLGAGALL 14

Db 6 FLGAGALL 14  
 RESULT 13  
 US-09-674-546A-3011  
 ; Sequence 3011, Application US/09674546A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Institute for Genomic Research  
 ; APPLICANT: Chiron Corporation  
 ; TITLE OF INVENTION: Neisseria Meningitidis Antigens and Compositions  
 ; FILE REFERENCE: CHIR-0334  
 ; CURRENT APPLICATION NUMBER: US/09/674,546A  
 ; CURRENT FILING DATE: 1999-04-30  
 ; NUMBER OF SEQ ID NOS: 3264  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 3011  
 ; LENGTH: 207  
 ; TYPE: PRT  
 ; ORGANISM: Neisseria meningitidis  
 ; US-09-674-546A-3011

Query Match 62.8%; Score 27; DB 5; Length 207;  
 Best Local Similarity 55.6%; Pred. No. 98;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLMAQAML 9  
 Db 6 FLGAGALL 14

RESULT 14  
 PCT-US03-10753-181  
 ; Sequence 181, Application PC/TUS0310753  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Shockley, Jay M.  
 ; APPLICANT: Schurr, Judy  
 ; TITLE OF INVENTION: Plant Acyl-CoA Synthetases  
 ; FILE REFERENCE: DOW-07655  
 ; CURRENT APPLICATION NUMBER: PCT/US03/10753  
 ; CURRENT FILING DATE: 2003-04-09  
 ; PRIOR APPLICATION NUMBER: 10/119,136  
 ; PRIOR FILING DATE: 2002-04-09  
 ; PRIOR APPLICATION NUMBER: 10/410,031  
 ; PRIOR FILING DATE: 2003-04-09  
 ; NUMBER OF SEQ ID NOS: 187  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 181  
 ; LENGTH: 212  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays  
 ; PCT-US03-10753-181

Query Match 62.8%; Score 27; DB 1; Length 212;  
 Best Local Similarity 83.3%; Pred. No. 1e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLMAQ 6  
 Db 65 FLAQ 70

RESULT 15  
 US-10-410-031-181  
 ; Sequence 181, Application US/10410031  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Shockley, Jay M.  
 ; APPLICANT: Schurr, Judy  
 ; APPLICANT: Brown, John A.  
 ; TITLE OF INVENTION: Plant Acyl-CoA Synthetases  
 ; FILE REFERENCE: DOW-07654  
 ; CURRENT APPLICATION NUMBER: US/10/410,031  
 ; CURRENT FILING DATE: 2003-04-09



```
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 181
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Zea mays
US-10-410-031-181
```

```
Query Match      62.8%; Score 27; DB 6; Length 212;
Best Local Similarity 83.3%; Pred No. 1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 FLMAQG 6
        ||:||||
Db       65 FLAQG 70
```

```
Search completed: October 7, 2003, 13:55:56
Job time : 3.8481 secs
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:04:57 ; Search time 16.5823 Seconds  
(without alignments)  
95.721 Million cell updates/sec

Title: US-09-807-512-25  
Perfect score: 45  
Sequence: 1 AMLAAQERRV 10  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues  
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_19Jun03:\*

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*

2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*

3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*

4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*

5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*

6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*

7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*

8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*

9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*

10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*

11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*

12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*

13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*

14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*

15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*

16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*

17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*

18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*

19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*

20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*

21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*

22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*

23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	100.0	10	21	AA1980
2	45	100.0	12	20	AA1981
3	45	100.0	13	20	AA1982
4	45	100.0	19	21	AA1983
5	45	100.0	30	23	AA1984
6	45	100.0	30	23	AA1985
7	45	100.0	58	20	AA1986
8	45	100.0	58	21	AA1987
9	45	100.0	58	23	AA1988

10	45	100.0	109	21	AA19854
11	45	100.0	3541	23	AAU85130
12	41	91.1	9	21	AAU70859
13	41	91.1	11	20	AAU05983
14	36	80.0	10	20	AAU05982
15	36	80.0	10	22	AA131332
16	36	80.0	10	23	ABG68803
17	32	71.1	9	20	AAU05981
18	32	71.1	198	23	ABP41770
19	32	71.1	456	22	AAU94081
20	32	71.1	673	20	AAU37106
21	32	71.1	586	23	ABU7446
22	32	71.1	772	23	ABU7543
23	32	71.1	1159	20	AAU32020
24	32	71.1	1159	21	AAU85405
25	32	71.1	1159	21	AAU85406
26	32	71.1	1159	22	AAU81812
27	32	71.1	1159	22	AAU70939
28	32	71.1	1159	23	AAU99167
29	32	71.1	1159	24	ABP71697
30	32	71.1	1159	24	ABP71255
31	32	71.1	1163	21	AAU83031
32	31	68.9	124	22	AAE01499
33	31	68.9	124	23	ABG63920
34	31	68.9	401	23	ABG91560
35	31	68.9	579	22	ABU60380
36	31	68.9	1419	23	ABU65081
37	30	66.7	78	22	ABG09298
38	30	66.7	194	23	ABP65687
39	30	66.7	227	22	ABU71097
40	30	66.7	429	22	ABU6886
41	30	66.7	429	24	ABP9266
42	30	66.7	549	22	ABG18981
43	30	66.7	561	23	ABU09587
44	30	66.7	758	22	AAU47227
45	30	66.7	758	24	ABU64938

ALIGNMENTS

RESULT 1

AAU70858  
ID AAU70858 standard; peptide; 10 AA.

XX AAU70858;

XX 31-JUL-2000 (first entry)

DT CAMEL16 immunogenic peptide of human CAMEL protein.

DE CAMEL; CTL-recognised Antigen on MELANOMA; cytotoxic T lymphocyte; CTL;  
KW tumour-associated antigen; LAGE-1; NY-ESO-1; anticancer; melanoma;  
KW human; cancer; immunotherapy; immunogenic peptide; immune response.

XX Homo sapiens.

XX WO200023584-A1.

XX 27-APR-2000.

XX 15-OCT-1999; 99WO-EP07832.

XX 16-OCT-1998; 98EP-0119583.

XX (BOEH ) BOEHRINGER INGELHEIM INT GMBH.  
XX (UYHO-) UNIV HOSPITAL LEIDEN.

XX Schrier PI, Aarnoudse CA, Heider K, Klade C;

XX WPI; 2000-339685/29.

XX Tumor-associated antigen useful for cancer immunotherapy is encoded by

PT the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA -  
 XX  
 PS Claim 7; Page 34; 73pp; English.  
 XX  
 CC The present sequence is an immunogenic peptide CAMEL 16, of the human  
 CC tumor-associated antigen CAMEL (Cytotoxic T lymphocytes (CTL)-recognised  
 CC Antigen on Melanoma). This peptide has the potential to bind to HLA-A2  
 CC and corresponds to residues 16-25 of the CAMEL protein. The CAMEL protein  
 CC is encoded by the LAGE-1 gene, a tumour-specific antigen. It is different  
 CC from the LAGE-1 protein, since it is translated from a different open  
 CC reading frame (ORF-1). It shows strong homology with NY-ESO-1, a melanoma  
 CC specific tumour antigen. The tumour-associated antigen displayed on  
 CC melanoma cells is recognised by cytotoxic T lymphocytes. CAMEL is  
 CC expressed in tumour cell lines, tumour tissues (e.g. breast and lung) and  
 CC in restricted number of healthy tissues. This sequence has anticancer  
 CC activity. CAMEL tumour antigen and immunogenic peptides derived from it  
 CC are useful for cancer immunotherapy. They have the potential to induce an  
 CC immune response, by eliciting a CTL response. The DNA molecule is used to  
 CC construct recombinant or fusion proteins.  
 XX  
 SQ Sequence 10 AA;  
 Query Match 100.0%; Score 45; DB 21; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.025;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AMLAAQERRV 10  
 DB 1 AMLAAQERRV 10  
 RESULT 2  
 AAY05984  
 ID AAY05984 standard; Peptide; 12 AA.  
 XX  
 AC AAY05984;  
 XX  
 DT 16-AUG-1999 (first entry)  
 XX  
 DE Human cancer antigen NY ESO-1/CAG-3 ORF2 cancer peptide.  
 XX  
 KW NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;  
 KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
 KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
 KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
 KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
 KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
 KW vaccine.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9918206-A2.  
 XX  
 PD 15-APR-1999.  
 XX  
 PF 21-SEP-1998; 98WO-US19609.  
 XX  
 PR 08-OCT-1997; 97US-0061428.  
 XX  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 PI Rosenberg SA, Wang RF;  
 XX  
 DR WPI; 1999-277270/23.  
 XX  
 PT Cancer antigen NY ESO1/CAG-3  
 XX  
 PS Disclosure; Page 12; 88pp; English.  
 CC  
 CC The present sequence represents a cancer peptide that corresponds  
 CC to amino acid residues 16-27 of human ESO-1/CAG-3 (or CAG-3) ORF2  
 CC (see AAY05966), a new and potent tumour antigen capable of eliciting  
 CC an antigen specific immune response by T cells. Cancer peptides  
 CC derived from CAG-3 ORF2, CAG-3 ORF1 (see AAY05965), portions of them  
 CC and their variants (see AAY05967-87), are useful as cancer vaccines  
 CC that protect against cancer. The invention provides: vectors and

CC derived from CAG-3 ORF2, CAG-3 ORF1 (see AAY05965), portions of them  
 CC and their variants (see AAY05967-87), are useful as cancer vaccines  
 CC that protect against cancer. The invention provides: vectors and  
 CC host cells (also useful as vaccines); a method of diagnosis of  
 CC cancer or precancer; a transgenic animal; antisense oligonucleotides  
 CC that inhibit expression of the cancer peptide or tumour antigen;  
 CC antibodies reacting with a CAG-3 cancer peptide, useful in  
 CC diagnostic and detection assays; and methods for preventing or  
 CC inhibiting cancer by administering a cancer peptide, with or without  
 CC an HLA molecule. The cancer peptides form part of, or are derived  
 CC from, cancers such as primary or metastatic melanoma, thymoma,  
 CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine  
 CC cancer, cervical cancer, bladder cancer, kidney cancer and  
 CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and  
 CC thyroid cancers. Melanoma is treated by inducing cancer-specific T  
 CC cells in vitro for subsequent return to a patient.  
 XX  
 SQ Sequence 12 AA;  
 Query Match 100.0%; Score 45; DB 20; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.03;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AMLAAQERRV 10  
 DB 1 AMLAAQERRV 10  
 RESULT 3  
 AAY05985  
 ID AAY05985 standard; Peptide; 13 AA.  
 XX  
 AC AAY05985;  
 XX  
 DT 16-AUG-1999 (first entry)  
 XX  
 DE Human cancer antigen NY ESO-1/CAG-3 ORF2 cancer peptide.  
 XX  
 KW NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;  
 KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
 KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
 KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
 KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
 KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
 KW vaccine.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9918206-A2.  
 XX  
 PD 15-APR-1999.  
 XX  
 PF 21-SEP-1998; 98WO-US19609.  
 XX  
 PR 08-OCT-1997; 97US-0061428.  
 XX  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 PI Rosenberg SA, Wang RF;  
 XX  
 DR WPI; 1999-277270/23.  
 XX  
 PT Cancer antigen NY ESO1/CAG-3  
 XX  
 PS Disclosure; Page 12; 88pp; English.  
 CC  
 CC The present sequence represents a cancer peptide that corresponds  
 CC to amino acid residues 15-27 of human ESO-1/CAG-3 (or CAG-3) ORF2  
 CC (see AAY05966), a new and potent tumour antigen capable of eliciting  
 CC an antigen specific immune response by T cells. Cancer peptides  
 CC derived from CAG-3 ORF2, CAG-3 ORF1 (see AAY05965), portions of them  
 CC and their variants (see AAY05967-87), are useful as cancer vaccines  
 CC that protect against cancer. The invention provides: vectors and

CC host cells (also useful as vaccines); a method of diagnosis of  
 CC cancer or precancer; a transgenic animal; antisense oligonucleotides  
 CC that inhibit expression of the cancer peptide or tumour antigen;  
 CC antibodies reacting with a CAG-3 cancer peptide, useful in  
 CC diagnostic and detection assays; and methods for preventing or  
 CC inhibiting cancer by administering a cancer peptide, with or without  
 CC an HLA molecule. The cancer peptides form part of, or are derived  
 CC from, cancers such as primary or metastatic melanoma, thymoma,  
 CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine  
 CC cancer, cervical cancer, bladder cancer, kidney cancer and  
 CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and  
 CC thyroid cancers. Melanoma is treated by inducing cancer-specific T  
 CC cells in vitro for subsequent return to a patient.  
 XX  
 SQ Sequence 13 AA;

Query Match 100.0%; Score 45; DB 20; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 0.033; 0; Indels 0; Gaps 0;  
 Matches 10; Conservative 0; Mismatches 0;

QY 1 AMLAAQERRV 10  
 |||||  
 Db 2 AMLAAQERRV 11

RESULT 4  
 AAY70904  
 ID AAY70904 standard; peptide; 19 AA.  
 AC AAY70904;  
 XX  
 XX 31-JUL-2000 (first entry)  
 XX Synthetic peptide, F4 from CAMEL protein.  
 XX  
 XX CAMEL; CTL-recognised Antigen on MELanoma; cytotoxic T lymphocyte; CTL;  
 KW tumour-associated antigen; LAGE-1; NY-ESO-1; anticancer; melanoma;  
 KW cancer; immunotherapy; immunogenic peptide; immune response; antibody.  
 XX  
 XX Homo sapiens.  
 OS Synthetic.  
 XX  
 XX WO200023584-A1.  
 XX  
 XX 27-APR-2000.  
 XX  
 XX 15-OCT-1999; 99WO-EP07832.  
 XX  
 XX 16-OCT-1998; 98EP-0119583.  
 XX  
 XX (BOEH ) BOEHRINGER INGELHEIM INT GMBH.  
 PA (UYHO-) UNIV HOSPITAL LEIDEN.  
 XX  
 XX Schrier PI, Aarnoudse CA, Heider K, Klade C;  
 XX  
 XX WPI; 2000-339685/29.  
 XX  
 XX Tumor-associated antigen useful for cancer immunotherapy is encoded by  
 PT the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA -  
 PT  
 XX Disclosure; Page 17; 73pp; English.  
 PS  
 XX The present sequence is the synthetic peptide F4, derived from the  
 CC hydrophobic regions of human CAMEL (Cytotoxic T lymphocytes (CTL)-  
 CC recognised Antigen on MELanoma), a tumour-associated antigen. It  
 CC corresponds to residues 15-31 of the CAMEL protein and is used to raise  
 CC antibodies. The Lys residue at both termini enable the peptides to be  
 CC linked to the carrier keyhole limpet haemocyanin (KLH) for immunisation.  
 CC CAMEL protein is encoded by the LAGE-1 gene, a tumour-specific antigen.  
 CC It shows strong homology with NY-ESO-1, a melanoma-specific tumour  
 CC antigen. The tumour-associated antigen displayed on melanoma cells is  
 CC recognised by cytotoxic T lymphocytes. CAMEL is expressed in tumour cell  
 CC lines, tumour tissues (e.g. breast and lung) and in restricted number of

CC healthy tissues. This sequence has anticancer activity. CAMEL tumour  
 CC antigen and immunogenic peptides derived from it are useful for cancer  
 CC immunotherapy. They have the potential to induce an immune response, by  
 CC eliciting a CTL response. The DNA molecule is used to construct  
 CC recombinant or fusion proteins.  
 XX  
 SQ Sequence 19 AA;

Query Match 100.0%; Score 45; DB 21; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 0.048; 0; Indels 0; Gaps 0;  
 Matches 10; Conservative 0; Mismatches 0;

QY 1 AMLAAQERRV 10  
 |||||  
 Db 3 AMLAAQERRV 12

RESULT 5  
 AAU85114  
 ID AAU85114 standard; Peptide; 30 AA.  
 XX  
 AC AAU85114;  
 XX  
 XX 08-MAY-2002 (first entry)  
 XX Human NYNSO1b segment 1.  
 XX  
 XX Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;  
 KW viral infection; human immunodeficiency virus; melanoma;  
 KW bacterial infection; Salmonella; Legionella; parasitic infection;  
 KW Trypanosoma; Toxoplasma; Giardia.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO200190197-A1.  
 PN  
 XX 29-NOV-2001.  
 PD  
 XX 25-MAY-2001; 2001WO-AU00622.  
 PF  
 XX 26-MAY-2000; 2000AU-0007761.  
 PR  
 XX (AUSU ) UNIV AUSTRALIAN NAT.  
 XX  
 XX Thomson SA, Ramshaw IA;  
 XX  
 XX WPI; 2002-147575/19.  
 DR N-PSDB; ABK36934.  
 XX  
 XX New synthetic polypeptides having several different segments of at  
 PT least one parent polypeptide linked together differently compared to  
 PT the linkage in the parent polypeptide, for inducing immune response  
 PT against a pathogen or cancer -  
 XX  
 XX Example 3; Fig 27; 364pp; English.  
 PS  
 XX The invention relates to a new synthetic polypeptide (I) comprising  
 CC several different segments of at least one parent polypeptide linked  
 CC together in a different relationship relative to their linkage in the  
 CC parent polypeptide to impede, abrogate or otherwise alter at least one  
 CC function associated with the parent polypeptide and for inducing an  
 CC immune response against a pathogen or cancer. Also included are a  
 CC synthetic polynucleotide encoding and a computer system for  
 CC designing the synthetic polypeptides. The synthetic polypeptides and  
 CC polynucleotides are referred to as a Savine. The synthetic polypeptide is  
 CC useful for modulating immune responses preferably directed against a  
 CC pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix,  
 CC colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone  
 CC liver, oesophagus, brain, testicle, uterus), as potentiating agents.  
 CC Compositions comprising the polypeptide may be used in the treatment or  
 CC prophylaxis against viral (such as infections caused by HIV (human  
 CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis  
 CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial

CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial  
 CC (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,  
 CC Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic  
 CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,  
 CC Trypanosoma, Toxoplasma and Giardia) infections. The present  
 CC sequence is a peptide derived from a parent protein used to  
 CC construct a vaccine of the invention.  
 CC  
 XX  
 SQ Sequence 30 AA;  
 Query Match 100.0%; Score 45; DB 23; Length 30;  
 Best Local Similarity 100.0%; Pred. No. 0.076;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 AMLAAQERRV 10  
 Db 3 AMLAAQERRV 12  
 RESULT 7  
 AAY05966  
 ID AAY05966 standard; Protein; 58 AA.  
 XX  
 AC AAY05966;  
 XX  
 DT 16-AUG-1999 (first entry)  
 XX  
 DE Human cancer antigen NY ESO-1/CAG-3 ORF2 protein.  
 XX  
 KW NY ESO-1/CAG-3 gene; CAG-3 gene; cancer antigen; human; leukaemia;  
 KW non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
 KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
 KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
 KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
 KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
 KW vaccine; ORF2.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9918206-A2.  
 XX  
 PD 15-APR-1999.  
 XX  
 PF 21-SEP-1998; 98WO-US19609.  
 XX  
 PR 08-OCT-1997; 97US-0061428.  
 XX  
 PR (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 PA Rosenberg SA, Wang RF;  
 PI WPI; 1999-277270/23.  
 XX  
 DR N-PSDB; AAX58599.  
 DR  
 XX  
 PT Cancer antigen NY ESO1/CAG-3  
 XX  
 PS Claim 5; Fig 3A; 88pp; English.  
 XX  
 CC The present sequence represents the ORF2 protein encoded by  
 CC open reading frame 2 of the human ESO-1/CAG-3 (or CAG-3) gene.  
 CC CAG-3 is a new and potent tumour antigen capable of eliciting an  
 CC antigen specific immune response by T cells. Cancer peptides  
 CC comprising ORF2, ORF1 (see AAY05965), portions of these peptides and  
 CC their variants (see AAY05967-87), are useful as cancer vaccines that  
 CC protect the recipient from development of cancer. The invention  
 CC provides: vectors and host cells (also useful as vaccines); a  
 CC method of diagnosis of cancer or precancer; a transgenic animal;  
 CC antisense oligonucleotides that inhibit expression of the cancer  
 CC peptide or tumour antigen; antibodies reacting with the CAG-3  
 CC cancer peptide, useful in diagnostic and detection assays; and  
 CC methods for preventing or inhibiting cancer by administering a  
 CC cancer peptide, with or without an HLA molecule. The cancer  
 CC peptides form part of, or are derived from, cancers such as

CC (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,  
 CC Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic  
 CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,  
 CC Trypanosoma, Toxoplasma and Giardia) infections. The present  
 CC sequence is a peptide derived from a parent protein used to  
 CC construct a vaccine of the invention.  
 CC  
 XX  
 SQ Sequence 30 AA;  
 Query Match 100.0%; Score 45; DB 23; Length 30;  
 Best Local Similarity 100.0%; Pred. No. 0.076;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 AMLAAQERRV 10  
 Db 18 AMLAAQERRV 27  
 RESULT 6  
 AAU85115  
 ID AAU85115 standard; Peptide; 30 AA.  
 XX  
 AC AAU85115;  
 XX  
 DT 08-MAY-2002 (first entry)  
 XX  
 DE Human NYSOLB segment 2.  
 XX  
 KW Vaccine; cancer; viral infection; HIV; hepatitis C virus;  
 KW viral infection; human immunodeficiency virus; melanoma;  
 KW bacterial infection; Salmonella; Legionella; parasitic infection;  
 KW Trypanosoma; Toxoplasma; Giardia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200190197-A1.  
 XX  
 PD 29-NOV-2001.  
 XX  
 PF 25-MAY-2001; 2001WO-AU00622.  
 XX  
 PR 26-MAY-2000; 2000AU-0007761.  
 XX  
 PR (AUSU ) UNIV AUSTRALIAN NAT.  
 XX  
 PI Thomson SA, Ramshaw IA;  
 XX  
 PI WPI; 2002-147575/19.  
 XX  
 DR N-PSDB; ABK36935.  
 DR  
 XX  
 PT New synthetic polypeptides having several different segments of at  
 PT least one parent polypeptide linked together differently compared to  
 PT the linkage in the parent polypeptide, for inducing immune response  
 PT against a pathogen or cancer -  
 XX  
 PS Example 3; Fig 27; 364pp; English.  
 XX  
 CC The invention relates to a new synthetic polypeptide (I) comprising  
 CC several different segments of at least one parent polypeptide linked  
 CC together in a different relationship relative to their linkage in the  
 CC parent polypeptide to impede, abrogate or otherwise alter at least one  
 CC function associated with the parent polypeptide and for inducing an  
 CC immune response against a pathogen or cancer. Also included are a  
 CC synthetic polynucleotide encoding and a computer system for  
 CC designing the synthetic polypeptides. The synthetic polypeptides and  
 CC polynucleotides are referred to as a vaccine. The synthetic polypeptide is  
 CC useful for modulating immune responses preferably directed against a  
 CC pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix,  
 CC colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone  
 CC liver, oesophagus, brain, testicle, uterus), as potentiating agents.  
 CC Compositions comprising the polypeptide may be used in the treatment or  
 CC prophylaxis against viral (such as infections caused by HIV (human  
 CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis

; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-2MFLB73091F09\_FLI.pep  
US-10-425-114A-71411

Query Match 73.2%; Score 30; DB 6; Length 317;  
Best Local Similarity 75.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAQERR 8  
|||:|:|  
Db 298 MLASQRR 305

## RESULT 12

US-10-425-114A-68789  
; Sequence 68789, Application US/10425114A  
; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114A

; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 68789

; LENGTH: 450  
; TYPE: PRT  
; ORGANISM: Zea mays

; FEATURE:  
; OTHER INFORMATION: Clone ID: 700342573\_FLI.pep  
US-10-425-114A-68789

Query Match 73.2%; Score 30; DB 6; Length 450;  
Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAQERR 8  
|||:|:|  
Db 433 MLASQRR 440

## RESULT 13

US-60-490-890-2340  
; Sequence 2340, Application US/60490890  
; GENERAL INFORMATION:

; APPLICANT: Li, Martha  
; APPLICANT: Rupnow, Brent A.  
; APPLICANT: Webster, Kevin R.  
; APPLICANT: Jackson, Donald  
; APPLICANT: Wong, Tai W.

; TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION  
; FILE REFERENCE: D0310 PSP

; CURRENT APPLICATION NUMBER: US/60/490,890  
; CURRENT FILING DATE: 2003-07-29

; NUMBER OF SEQ ID NOS: 2779  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 2340

; LENGTH: 267  
; TYPE: PRT  
; ORGANISM: Homo sapiens

US-60-490-890-2340

Query Match 70.7%; Score 29; DB 7; Length 267;  
Best Local Similarity 85.7%; Pred. No. 1.6e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAQERR 9  
||:|:|  
Db 34 AAERR 40

## RESULT 14

US-10-425-114A-49793.  
; Sequence 49793, Application US/10425114A  
; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114A

; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 49793

; LENGTH: 274  
; TYPE: PRT  
; ORGANISM: Zea mays

; FEATURE:  
; OTHER INFORMATION: Clone ID: 700469566\_FLI.pep  
US-10-425-114A-49793

Query Match 70.7%; Score 29; DB 6; Length 274;  
Best Local Similarity 85.7%; Pred. No. 1.6e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAAQERR 8  
||:|:|  
Db 268 LASQERR 274

## RESULT 15

US-10-425-114A-70887  
; Sequence 70887, Application US/10425114A  
; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114A

; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 70887

; LENGTH: 407  
; TYPE: PRT  
; ORGANISM: Zea mays

; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3354-066-A4\_FLI.pep  
US-10-425-114A-70887

Query Match 70.7%; Score 29; DB 6; Length 407;  
Best Local Similarity 75.0%; Pred. No. 2.1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAAQERR 8  
:|:|:|  
Db 232 LAAQERR 239

Search completed: October 7, 2003, 13:55:57  
Job time : 2.8481 secs

```

, TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
,
, FILE REFERENCE: 38-21(53313)B
, CURRENT APPLICATION NUMBER: US/10/425,114A
, CURRENT FILING DATE: 2003-04-28
, NUMBER OF SEQ ID NOS: 73128
,
, SEQ ID NO 67943
, LENGTH: 164
,

```

```

; FEATURE:
; OTHER INFORMATION: Clone ID: UC-2MFLMO17186F05_FLI.pcp
US-10-425-114A-67943

```

Query Match 75.6%; Score 31; DB 6; Length 164;  
Best Local Similarity 87.5%; Pred. No. 53;  
Matches 7: Conservative 0; Mismatches 1; Indels

Qy 2 LAQERRV 9  
| | | | |  
Db 150 LEAQERRV 157

RESULT 10  
PCT-US03-24669-286

APPLICANT: Macina, Roberto  
APPLICANT: Salceda, Susana  
APPLICANT: Liu, Chenghua  
APPLICANT: Sun, Yongming  
APPLICANT: Turner, Leah

GENERAL INFORMATION:  
APPLICANT: diaDexus, Inc.  
Sequence 286, Application P

```

; TITLE OF INVENTION: Compositions and Methods
; FILE REFERENCE: DEX-0443
; CURRENT APPLICATION NUMBER: PCT/US03/246669

```

; CURRENT FILING DATE: 2003-08-06  
 ;  
 ; PRIOR APPLICATION NUMBER: US 60/401,469  
 ;  
 ; PRIOR FILING DATE: 2002-08-06  
 ;

```

; NUMBER OF SEQ ID NOS: 396
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 286
; LENGTH: 347

```

```

; TYPE: PRT
; ORGANISM: Homo sapien
PCT-IIS03-24669-286

```

Query Match 75.6%; Score 31; DB 1; Length 347;  
Best Local Similarity 87.5%; Pred. No. 88;  
Matches 7; Conservative 0; Mismatches 1; Indels

QY 1 MLAAQERR 8  
||| |||  
Db 108 MLANOERR 115

RESULT 11  
US-10-425-114A-71411

APPLICANT: Liu, Jingdong  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K  
APPLICANT: Screen, Steven E  
APPLICANT: Tabaska, Jack E  
APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules  
 TITLE OF INVENTION: Plants and Uses Thereof  
 FILE REFERENCE: 38-21(53313)B  
 CURRENT APPLICATION NUMBER: US/10/425,114A  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 73128  
 SEQ ID NO 71411  
 LENGTH: 317

APPLICANT: Liu, Jingdong  
 APPLICANT: Zhou, Yihua  
 APPLICANT: Kovalic, David K.  
 APPLICANT: Screen, Steven E.  
 APPLICANT: Tabaska, Jack E.  
 APPLICANT: Cao, Yongwei  
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 and/or Therapeutic for Plant Improvement

FILE REFERENCE: 50 21(5515)D  
CURRENT APPLICATION NUMBER: US/10/425,114A  
CURRENT FILING DATE: 2003-04-28

```

; SEQ ID NO 70170
;
; LENGTH: 736
; TYPE: PRT
;

```

```

; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73120C10_FLI.pcp
; NS-10-425-114A-70170

```

Query Match	78.0%	Score 32;	DB 6;	Length 736;
Best Local Similarity	100.0%;	Pred. No. 99;		
Matches	7;	Conservative 0;	Mismatches 0;	Indels 0;
Gaps	0;			

Qy 3 AAQRRV 9  
|||||  
Db 568 AAOERRV 574

RESULT 8  
US-10-425-114A-46933  
. sequence 46933. Application US/10425114A

; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jingdong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovalic, David K  
 ; APPLICANT: Screen, Steven E  
 ; APPLICANT: Tabaska, Jack E

**APPLICANT:** Cao, Yongwei  
**TITLE OF INVENTION:** Nucleic Acid Molecules and Other Molecules Associated with  
the Growth and Development of Plants and Uses Thereof for Plant Improvement

```

; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128

```

```

; SEQ ID NO 46933
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:

```

```

; FEATURE:
; OTHER INFORMATION: Clone ID: 700620710_FLI.pep
US-10-425-114A-46933

```

Query Match 75.6%; Score 31; DB 6; Length 162;  
Best Local Similarity 87.5%; Pred. No. 53;  
Matches 7; Conservative 0; Mismatches 1; Indels

QY 2 LAAQRRV 9  
| | | | |  
Dp 148 LEAQRRV 155

RESULT 9  
US-10-425-114A-67943  
: Sequence 67943. Application US/10425114A

; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jingdong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Screen, Steven E  
 ; APPLICANT: Tabaska, Jack E  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic A

APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With



```
; FEATURE:
; OTHER INFORMATION: NYSolb segment 2
US-10-296-734-1424

Query Match      100.0%; Score 41; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.36;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAAQERRV 9
Db 4 MLAAQERRV 12

RESULT 3
US-10-296-734-833
; Sequence 833, Application US/10296734
; GENERAL INFORMATION:
; APPLICANT: Thompson, Scott A
; APPLICANT: Ramshaw, Ian A
; TITLE OF INVENTION: Synthetic molecules and uses therefor
; FILE REFERENCE: Savine
; CURRENT APPLICATION NUMBER: US/10/296,734
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: AU P07761/00
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 833
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: NYSolb consensus polypeptide
US-10-296-734-833

Query Match      100.0%; Score 41; DB 6; Length 58;
Best Local Similarity 100.0%; Pred. No. 0.56;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAAQERRV 9
Db 17 MLAAQERRV 25

RESULT 4
US-10-296-734-1454
; Sequence 1454, Application US/10296734
; GENERAL INFORMATION:
; APPLICANT: Thompson, Scott A
; APPLICANT: Ramshaw, Ian A
; TITLE OF INVENTION: Synthetic molecules and uses therefor
; FILE REFERENCE: Savine
; CURRENT APPLICATION NUMBER: US/10/296,734
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: AU P07761/00
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1454
; LENGTH: 3541
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Melanoma cancer specific savine
US-10-296-734-1454

Query Match      100.0%; Score 41; DB 6; Length 3541;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAAQERRV 9
Db 184 MLAAQERRV 192

RESULT 5
US-10-654-102-35
; Sequence 35, Application US/10654102
; GENERAL INFORMATION:
; APPLICANT: CHAN, LAWRENCE
; APPLICANT: KOJIMA, HIDEOTO
; TITLE OF INVENTION: INDUCTION OF PANCREATIC ISLET FORMATION
; FILE REFERENCE: P02409US1
; CURRENT APPLICATION NUMBER: US/10/654,102
; CURRENT FILING DATE: 2003-09-03
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (22)..(94)
; OTHER INFORMATION: x = anything
US-10-654-102-35

Query Match      78.0%; Score 32; DB 6; Length 103;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAQERRV 9
Db 31 AAQERRV 37

RESULT 6
US-10-425-114A-64836
; Sequence 64836, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 64836
; LENGTH: 446
; TYPE: PRT
; ORGANISM: zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4720-012-D11_FLI.pep
US-10-425-114A-64836

Query Match      78.0%; Score 32; DB 6; Length 446;
Best Local Similarity 66.7%; Pred. No. 71;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAAQERRV 9
Db 65 LLAAQRRRL 73

RESULT 7
US-10-425-114A-70170
; Sequence 70170, Application US/10425114A
; GENERAL INFORMATION:
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:26:33 ; Search time 2.8481 Seconds  
(without alignments)  
92.039 Million cell updates/sec

Title: US-09-807-512-26  
Perfect score: 41  
Sequence: 1 MLAAQERRV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 136899 seqs, 29126274 residues

Total number of hits satisfying chosen parameters: 136899

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database :
- 1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep.\*
  - 2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep.\*
  - 3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep.\*
  - 4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep.\*
  - 5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep.\*
  - 6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep.\*
  - 7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	41	100.0	30	6	US-10-296-734-1422
2	41	100.0	30	6	US-10-296-734-1424
3	41	100.0	58	6	US-10-296-734-833
4	41	100.0	3541	6	US-10-296-734-1454
5	32	78.0	103	6	US-10-654-102-35
6	32	78.0	446	6	US-10-425-114A-64836
7	32	78.0	736	6	US-10-425-114A-70170
8	31	75.6	162	6	US-10-425-114A-46933
9	31	75.6	164	6	US-10-425-114A-67943
10	31	75.6	347	1	PCT-US03-24669-286
11	30	73.2	317	6	US-10-425-114A-71411
12	30	73.2	450	6	US-10-425-114A-68789
13	29	70.7	267	7	US-60-490-890-2340
14	29	70.7	274	6	US-10-425-114A-49793
15	29	70.7	407	6	US-10-425-114A-70887
16	29	70.7	504	6	US-10-425-114A-70248
17	29	70.7	794	1	PCT-US03-26780-2438
18	28	68.3	165	6	US-10-463-980-2
19	28	68.3	166	5	US-09-994-425-1
20	28	68.3	166	5	US-09-994-425-13
21	28	68.3	188	6	US-10-652-995-1
22	28	68.3	188	6	US-10-653-350-1
23	28	68.3	216	6	US-10-425-114A-71626
24	28	68.3	250	6	US-10-425-114A-53824
25	28	68.3	264	6	US-10-425-114A-62722
26	28	68.3	284	6	US-10-425-114A-46674

27	28	68.3	291	6	US-10-425-114A-71228	Sequence 71228, A
28	28	68.3	330	6	US-10-425-114A-69837	Sequence 69837, A
29	28	68.3	354	1	PCT-US03-26780-2245	Sequence 2245, Ap
30	28	68.3	424	6	US-10-425-114A-67531	Sequence 67531, A
31	28	68.3	424	6	US-10-425-114A-70829	Sequence 70829, A
32	28	68.3	433	6	US-10-425-114A-57335	Sequence 57335, A
33	28	68.3	441	6	US-10-425-114A-64527	Sequence 64527, A
34	28	68.3	442	6	US-10-425-114A-43237	Sequence 43237, A
35	28	68.3	451	6	US-10-425-114A-46249	Sequence 46249, A
36	28	68.3	489	6	US-10-425-114A-49473	Sequence 49473, A
37	28	68.3	506	1	PCT-US03-28227-3700	Sequence 3700, Ap
38	28	68.3	509	1	PCT-US03-28227-3698	Sequence 3698, Ap
39	28	68.3	509	1	PCT-US03-28227-3699	Sequence 3699, Ap
40	28	68.3	525	1	PCT-US03-28227-3697	Sequence 3697, Ap
41	28	68.3	573	7	US-60-500-337-2423	Sequence 2423, Ap
42	28	68.3	581	6	US-10-425-114A-62613	Sequence 62613, A
43	28	68.3	590	6	US-10-425-114A-69685	Sequence 69685, A
44	28	68.3	621	1	PCT-US03-14382-446	Sequence 446, App
45	28	68.3	621	1	PCT-US03-14382A-446	Sequence 446, App

ALIGNMENTS

RESULT 1  
US-10-296-734-1422  
; Sequence 1422, Application US/10296734  
; GENERAL INFORMATION:  
; APPLICANT: Thompson, Scott A  
; APPLICANT: Ramshaw, Ian A  
; TITLE OF INVENTION: Synthetic molecules and uses therefor  
; FILE REFERENCE: Savine  
; CURRENT APPLICATION NUMBER: US/10/296,734  
; CURRENT FILING DATE: 2003-08-04  
; PRIOR APPLICATION NUMBER: AU PQ7761/00  
; PRIOR FILING DATE: 2000-05-26  
; NUMBER OF SEQ ID NOS: 1507  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1422  
; LENGTH: 30  
; TYPE: PRT  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: NYSolb segment 1  
US-10-296-734-1422

Query Match 100.0%; Score 41; DB 6; Length 30;  
Best Local Similarity 100.0%; Pred. No. 0.36;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAAQERRV 9  
Db 19 MLAAQERRV 27  
|||||

RESULT 2  
US-10-296-734-1424  
; Sequence 1424, Application US/10296734  
; GENERAL INFORMATION:  
; APPLICANT: Thompson, Scott A  
; APPLICANT: Ramshaw, Ian A  
; TITLE OF INVENTION: Synthetic molecules and uses therefor  
; FILE REFERENCE: Savine  
; CURRENT APPLICATION NUMBER: US/10/296,734  
; CURRENT FILING DATE: 2003-08-04  
; PRIOR APPLICATION NUMBER: AU PQ7761/00  
; PRIOR FILING DATE: 2000-05-26  
; NUMBER OF SEQ ID NOS: 1507  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1424  
; LENGTH: 30  
; TYPE: PRT  
; ORGANISM: Artificial

**THIS PAGE BLANK (USPTO)**

Query Match 100.0%; Score 41; DB 19; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.43;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAGERRV 9  
      |||||  
Db 17 MLAGERRV 25

Search completed: October 7, 2003, 13:55:01  
Job time : 70.8608 secs

```
RESULT 11
US-09-529-206-50
; Sequence 50, Application US/09529206
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America as
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; FILE REFERENCE: 20264269PC
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206-50

Query Match      100.0%; Score 41; DB 19; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAAQERRV 9
Db 3 MLAAQERRV 11

RESULT 12
US-09-529-206A-50
; Sequence 50, Application US/09529206A
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; FILE REFERENCE: 20264269US1
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: PCT/US98/19609
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 50
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206A-50

Query Match      100.0%; Score 41; DB 19; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAAQERRV 9
Db 3 MLAAQERRV 11

RESULT 13
US-09-529-206B-50
; Sequence 50, Application US/09529206B
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; FILE REFERENCE: 20264269US1
; CURRENT APPLICATION NUMBER: US/09/529,206B
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206B-50

Query Match      100.0%; Score 41; DB 19; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAAQERRV 9
Db 3 MLAAQERRV 11

RESULT 14
US-09-529-206D-50
; Sequence 50, Application US/09529206D
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; FILE REFERENCE: 218791
; CURRENT APPLICATION NUMBER: US/09/529,206D
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: PCT/US98/19609
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 50
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206D-50

Query Match      100.0%; Score 41; DB 19; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAAQERRV 9
Db 3 MLAAQERRV 11

RESULT 15
US-09-529-206-56
; Sequence 56, Application US/09529206
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America as
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; FILE REFERENCE: 20264269PC
; CURRENT APPLICATION NUMBER: US/09/529,206
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 56
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206-56

Query Match      100.0%; Score 41; DB 19; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAAQERRV 9
Db 3 MLAAQERRV 11
```

; CURRENT APPLICATION NUMBER: US/09/529,206D  
; CURRENT FILING DATE: 2000-06-13  
; PRIOR APPLICATION NUMBER: PCT/US98/19609  
; PRIOR FILING DATE: 1998-09-21  
; PRIOR APPLICATION NUMBER: US60/061,428  
; PRIOR FILING DATE: 1997-10-08  
; NUMBER OF SEQ ID NOS: 127  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 48  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-206D-48

Query Match 100.0%; Score 41; DB 19; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.16;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAQERRV 9  
| | | | | | | | | |  
DB 1 MAAQERRV 9

## RESULT 7

US-09-529-206-49  
; Sequence 49, Application US/09529206  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America as  
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene  
; FILE REFERENCE: 20264269PC  
; CURRENT APPLICATION NUMBER: US/09/529,206  
; CURRENT FILING DATE: 2000-06-13  
; PRIOR APPLICATION NUMBER: US60/061,428  
; PRIOR FILING DATE: 1997-10-08  
; NUMBER OF SEQ ID NOS: 106  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 49  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-206-49

Query Match 100.0%; Score 41; DB 19; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.18;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAQERRV 9  
| | | | | | | | | |  
DB 2 MAAQERRV 10

## RESULT 8

US-09-529-206A-49  
; Sequence 49, Application US/09529206A  
; GENERAL INFORMATION:  
; APPLICANT: Wang Rong, Fu  
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene  
; FILE REFERENCE: 20264269US1  
; CURRENT APPLICATION NUMBER: US/09/529,206A  
; CURRENT FILING DATE: 2000-06-13  
; PRIOR APPLICATION NUMBER: PCT/US98/19609  
; PRIOR FILING DATE: 1998-09-21  
; PRIOR APPLICATION NUMBER: US60/061,428  
; PRIOR FILING DATE: 1997-10-08  
; NUMBER OF SEQ ID NOS: 127  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 49  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Homo sapiens

## US-09-529-206A-49

Query Match 100.0%; Score 41; DB 19; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.18;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAQERRV 9  
| | | | | | | | | |  
DB 2 MAAQERRV 10

## RESULT 9

US-09-529-206B-49  
; Sequence 49, Application US/09529206B  
; GENERAL INFORMATION:  
; APPLICANT: Wang Rong, Fu  
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene  
; FILE REFERENCE: 20264269US1  
; CURRENT APPLICATION NUMBER: US/09/529,206B  
; CURRENT FILING DATE: 2000-06-13  
; PRIOR APPLICATION NUMBER: PCT/US98/19609  
; PRIOR FILING DATE: 1998-09-21  
; PRIOR APPLICATION NUMBER: US60/061,428  
; PRIOR FILING DATE: 1997-10-08  
; NUMBER OF SEQ ID NOS: 127  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 49  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-206B-49

Query Match 100.0%; Score 41; DB 19; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.18;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAQERRV 9  
| | | | | | | | | |  
DB 2 MAAQERRV 10

## RESULT 10

US-09-529-206D-49  
; Sequence 49, Application US/09529206D  
; GENERAL INFORMATION:  
; APPLICANT: Wang Rong, Fu  
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene  
; FILE REFERENCE: 218791  
; CURRENT APPLICATION NUMBER: US/09/529,206D  
; CURRENT FILING DATE: 2000-06-13  
; PRIOR APPLICATION NUMBER: PCT/US98/19609  
; PRIOR FILING DATE: 1998-09-21  
; PRIOR APPLICATION NUMBER: US60/061,428  
; PRIOR FILING DATE: 1997-10-08  
; NUMBER OF SEQ ID NOS: 127  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 49  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-206D-49

Query Match 100.0%; Score 41; DB 19; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.18;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAQERRV 9  
| | | | | | | | | |  
DB 2 MAAQERRV 10

Query Match 100.0%; Score 41; DB 23; Length 9;  
Best Local Similarity 100.0%; Pred. No. 5.3e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAAQERRV 9  
Db 1 MLAAQERRV 9

## RESULT 2

US-09-807-512-25  
; Sequence 25, Application US/09807512  
; GENERAL INFORMATION:  
; APPLICANT: Schrier, Peter I.  
; APPLICANT: Aarnoudse, Corlien  
; APPLICANT: Heider, Karl-Heinz  
; APPLICANT: Klade, Christoph  
; TITLE OF INVENTION: Camel, An Alternative Translation Product of the Tumor  
; TITLE OF INVENTION: Antigen-Lage 1  
; FILE REFERENCE: 0652.2200000  
; CURRENT APPLICATION NUMBER: US/09/807,512  
; PRIOR FILING DATE: 2001-04-16  
; PRIOR APPLICATION NUMBER: PCT/EP99/07832  
; PRIOR FILING DATE: 1999-10-15  
; PRIOR APPLICATION NUMBER: EP 98119583.7  
; PRIOR FILING DATE: 1998-10-16  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 25  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-807-512-25

Query Match 100.0%; Score 41; DB 23; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.15;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAAQERRV 9  
Db 2 MLAAQERRV 10

## RESULT 3

US-09-529-206-48  
; Sequence 48, Application US/09529206  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America as  
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene  
; FILE REFERENCE: 20264269PC  
; CURRENT APPLICATION NUMBER: US/09/529,206  
; PRIOR FILING DATE: 2000-06-13  
; PRIOR APPLICATION NUMBER: US60/061,428  
; PRIOR FILING DATE: 1997-10-08  
; NUMBER OF SEQ ID NOS: 106  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 48  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-206-48

Query Match 100.0%; Score 41; DB 19; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.16;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAAQERRV 9  
Db 1 MLAAQERRV 9

## RESULT 4

US-09-529-206A-48  
; Sequence 48, Application US/09529206A  
; GENERAL INFORMATION:  
; APPLICANT: Wang Rong, Fu  
; APPLICANT: Rosenberg, Steven  
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene  
; TITLE OF INVENTION: Encoding Same  
; FILE REFERENCE: 20264269US1  
; CURRENT APPLICATION NUMBER: US/09/529,206A  
; CURRENT FILING DATE: 2000-06-13  
; PRIOR APPLICATION NUMBER: PCT/US98/19609  
; PRIOR FILING DATE: 1998-09-21  
; PRIOR APPLICATION NUMBER: US60/061,428  
; PRIOR FILING DATE: 1997-10-08  
; NUMBER OF SEQ ID NOS: 127  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 48  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-206A-48

Query Match 100.0%; Score 41; DB 19; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.16;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAAQERRV 9  
Db 1 MLAAQERRV 9

## RESULT 5

US-09-529-206B-48  
; Sequence 48, Application US/09529206B  
; GENERAL INFORMATION:  
; APPLICANT: Wang Rong, Fu  
; APPLICANT: Rosenberg, Steven  
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene  
; TITLE OF INVENTION: Encoding Same  
; FILE REFERENCE: 20264269US1  
; CURRENT APPLICATION NUMBER: US/09/529,206B  
; CURRENT FILING DATE: 2000-06-13  
; PRIOR APPLICATION NUMBER: PCT/US98/19609  
; PRIOR FILING DATE: 1998-09-21  
; PRIOR APPLICATION NUMBER: US60/061,428  
; PRIOR FILING DATE: 1997-10-08  
; NUMBER OF SEQ ID NOS: 127  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 48  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-206B-48

Query Match 100.0%; Score 41; DB 19; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.16;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAAQERRV 9  
Db 1 MLAAQERRV 9

## RESULT 6

US-09-529-206D-48  
; Sequence 48, Application US/09529206D  
; GENERAL INFORMATION:  
; APPLICANT: Wang Rong, Fu  
; APPLICANT: Rosenberg, Steven  
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene  
; TITLE OF INVENTION: Encoding Same  
; FILE REFERENCE: 218791

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model  
Run on: October 7, 2003, 13:25:43 ; Search time 70.8608 Seconds  
(without alignments)  
115.568 Million cell updates/sec

Title: US-09-807-512-26  
Perfect score: 41  
Sequence: 1 MLAGERRV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues  
Total number of hits satisfying chosen parameters: 5728757

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main:\*

1: /cgn2\_6/ptodata/1/paa/pctus\_COMB.pcp.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pcp.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pcp.\*  
4: /cgn2\_6/ptodata/1/paa/US080\_COMB.pcp.\*  
5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pcp.\*  
6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pcp.\*  
7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pcp.\*  
8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pcp.\*  
9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pcp.\*  
10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pcp.\*  
11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pcp.\*  
12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pcp.\*  
13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pcp.\*  
14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pcp.\*  
15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pcp.\*  
16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pcp.\*  
17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pcp.\*  
18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pcp.\*  
19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pcp.\*  
20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pcp.\*  
21: /cgn2\_6/ptodata/1/paa/US097A\_COMB.pcp.\*  
22: /cgn2\_6/ptodata/1/paa/US097B\_COMB.pcp.\*  
23: /cgn2\_6/ptodata/1/paa/US098\_COMB.pcp.\*  
24: /cgn2\_6/ptodata/1/paa/US099A\_COMB.pcp.\*  
25: /cgn2\_6/ptodata/1/paa/US099B\_COMB.pcp.\*  
26: /cgn2\_6/ptodata/1/paa/US100\_COMB.pcp.\*  
27: /cgn2\_6/ptodata/1/paa/US101\_COMB.pcp.\*  
28: /cgn2\_6/ptodata/1/paa/US102\_COMB.pcp.\*  
29: /cgn2\_6/ptodata/1/paa/US103\_COMB.pcp.\*  
30: /cgn2\_6/ptodata/1/paa/US104\_COMB.pcp.\*  
31: /cgn2\_6/ptodata/1/paa/US106\_COMB.pcp.\*  
32: /cgn2\_6/ptodata/1/paa/US60\_COMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	9	23	US-09-807-512-26 Sequence 26, Appl

2	41	100.0	10	23	US-09-807-512-25	Sequence 25, Appl
3	41	100.0	11	19	US-09-529-206-48	Sequence 48, Appl
4	41	100.0	11	19	US-09-529-206A-48	Sequence 48, Appl
5	41	100.0	11	19	US-09-529-206B-48	Sequence 48, Appl
6	41	100.0	11	19	US-09-529-206D-48	Sequence 48, Appl
7	41	100.0	12	19	US-09-529-206-49	Sequence 49, Appl
8	41	100.0	12	19	US-09-529-206A-49	Sequence 49, Appl
9	41	100.0	12	19	US-09-529-206B-49	Sequence 49, Appl
10	41	100.0	12	19	US-09-529-206D-49	Sequence 49, Appl
11	41	100.0	13	19	US-09-529-206A-50	Sequence 50, Appl
12	41	100.0	13	19	US-09-529-206B-50	Sequence 50, Appl
13	41	100.0	13	19	US-09-529-206D-50	Sequence 50, Appl
14	41	100.0	25	19	US-09-529-206-56	Sequence 56, Appl
15	41	100.0	25	19	US-09-529-206A-56	Sequence 56, Appl
16	41	100.0	25	19	US-09-529-206B-56	Sequence 56, Appl
17	41	100.0	25	19	US-09-529-206D-56	Sequence 56, Appl
18	41	100.0	58	19	US-09-529-206-5	Sequence 5, Appl
19	41	100.0	58	19	US-09-529-206A-5	Sequence 5, Appl
20	41	100.0	58	19	US-09-529-206B-5	Sequence 5, Appl
21	41	100.0	58	19	US-09-529-206D-5	Sequence 5, Appl
22	41	100.0	109	23	US-09-807-512-10	Sequence 10, Appl
23	41	100.0	109	23	US-09-807-512-2	Sequence 2, Appl
24	41	100.0	10	17	US-09-336-091-45	Sequence 45, Appl
25	36	87.8	10	19	US-09-529-206-47	Sequence 47, Appl
26	36	87.8	10	19	US-09-529-206A-47	Sequence 47, Appl
27	36	87.8	10	19	US-09-529-206B-47	Sequence 47, Appl
28	36	87.8	10	19	US-09-529-206D-47	Sequence 47, Appl
29	36	87.8	325	27	US-10-155-881-8908	Sequence 8908, Ap
30	36	87.8	182	18	US-09-417-507-44277	Sequence 44277, A
31	35	85.4	341	18	US-09-428-944-1639	Sequence 1639, Ap
32	34	82.9	350	18	US-09-417-507-29765	Sequence 29765, A
33	34	82.9	30	30	US-10-437-963-182418	Sequence 182418, A
34	33	80.5	129	30	US-10-382-122A-66900	Sequence 66900, A
35	33	80.5	482	28	US-10-382-122A-66900	Sequence 122354, A
36	33	80.5	1429	22	US-09-791-537-122354	Sequence 5819, Ap
37	33	80.5	1429	22	US-10-369-493-5819	Sequence 5819, Ap
38	33	80.5	1429	32	US-60-360-039-5819	Sequence 46, Appl
39	32	78.0	9	19	US-09-529-206-46	Sequence 46, Appl
40	32	78.0	9	19	US-09-529-206A-46	Sequence 46, Appl
41	32	78.0	9	19	US-09-529-206B-46	Sequence 123, App
42	32	78.0	9	19	US-09-529-206D-46	Sequence 123, App
43	32	78.0	9	19	US-09-529-206B-123	Sequence 123, App
44	32	78.0	9	19	US-09-529-206D-123	Sequence 123, App
45	32	78.0	9	19	US-09-529-206D-123	Sequence 123, App

ALIGNMENTS

RESULT 1  
US-09-807-512-26  
Sequence 26, Application US/09807512  
GENERAL INFORMATION:  
APPLICANT: Schrier, Peter I.  
APPLICANT: Aarnoudse, Corlien  
APPLICANT: Heider, Karl-Heinz  
APPLICANT: Klade, Christoph  
TITLE OF INVENTION: Camel, An Alternative Translation Product of the Tumor  
FILE REFERENCE: 0652.2200000  
CURRENT APPLICATION NUMBER: US/09/807,512  
CURRENT FILING DATE: 2001-04-16  
PRIOR APPLICATION NUMBER: PCT/EP99/07832  
PRIOR FILING DATE: 1999-10-15  
PRIOR APPLICATION NUMBER: EP 98119583.7  
PRIOR FILING DATE: 1998-10-16  
NUMBER OF SEQ ID NOS: 28  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 26  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-807-512-26



**THIS PAGE BLANK (USPTO)**

US-10-156-761-10790  
; Sequence 10790, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; PRIOR FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 10790  
; LENGTH: 446  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-10790

Query Match 70.7%; Score 29; DB 15; Length 446;  
Best Local Similarity 75.0%; Pred. No. 4.9e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LAAQERRV 9  
||| |||  
Db 42 LAAARRRI 49

US-10-156-761-12826  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; PRIOR FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 12826  
; LENGTH: 636  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-12826

Query Match 70.7%; Score 29; DB 15; Length 636;  
Best Local Similarity 85.7%; Pred. No. 7e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 AAQERRV 9  
||| |||  
Db 311 AAQERRM 317

Search completed: October 7, 2003, 13:58:58  
Job time : 10.7975 secs

US-09-955-999-110  
; Sequence 110, Application US/09955999  
; Publication No. US20030036505A1  
; GENERAL INFORMATION:  
; APPLICANT: Barash et al.  
; TITLE OF INVENTION: Signal Transduction Pathway Component Polynucleotides, Polypeptides, and Methods Based Thereon  
; FILE REFERENCE: PT086P1  
; CURRENT APPLICATION NUMBER: US/09/955,999  
; CURRENT FILING DATE: 2001-09-20  
; PRIOR APPLICATION NUMBER: 60/234,997  
; PRIOR FILING DATE: 2000-09-25  
; NUMBER OF SEQ ID NOS: 139  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 110  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-955-999-110

Query Match 70.7%; Score 29; DB 11; Length 455;  
Best Local Similarity 55.6%; Pred. No. 5e+02;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAQERRV 9  
||| : |||  
Db 438 MAAARKRI 446

RESULT 14

US-10-156-761-12826  
; Sequence 12826, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN

RESULT 15

; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 12440  
; LENGTH: 292  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-09-815-242-12440

Query Match 70.7%; Score 29; DB 9; Length 292;  
Best Local Similarity 66.7%; Pred. No. 3.2e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MIAAQQRRV 9  
DB 57 MIAAQQRRV 65

## RESULT 10

US-10-156-761-7637  
; Sequence 7637, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 7637  
; LENGTH: 331  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-7637

Query Match 70.7%; Score 29; DB 15; Length 331;  
Best Local Similarity 75.0%; Pred. No. 3.6e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAAQRRV 9  
DB 83 LAAQRRV 90

## RESULT 11

US-09-817-513A-2  
; Sequence 2, Application US/09817513A  
; Publication No. US20030044958A1  
; GENERAL INFORMATION:  
; APPLICANT: Staunton, Donald E.  
; APPLICANT: Lipsky, Brian P.  
; TITLE OF INVENTION: Cytoplasmic Modulators of Integrin  
Binding/Signaling

NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker, 6300 Sears Tower  
CITY: Chicago  
STATE: Illinois

COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/817,513A  
FILING DATE: 28-Jun-2001  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Young J. Suh  
REGISTRATION NUMBER: P-41,337  
REFERENCE/DOCKET NUMBER: 27866/33886  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
TELEFAX: 312-474-0448  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 399 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-817-513A-2

Query Match 70.7%; Score 29; DB 11; Length 399;  
Best Local Similarity 55.6%; Pred. No. 4.4e+02;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIAAQQRRV 9  
DB 382 MIAAQQRRV 390

## RESULT 12

US-10-156-761-8533  
; Sequence 8533, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 8533  
; LENGTH: 431  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-8533

Query Match 70.7%; Score 29; DB 15; Length 431;  
Best Local Similarity 66.7%; Pred. No. 4.7e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MIAAQQRRV 9  
DB 356 MIAAQQRRV 364

## RESULT 13

```
; Sequence 4027, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4027
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (44)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (57)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (61)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (80)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-764-891-4027

Query Match          70.7%; Score 29; DB 11; Length 153;
Best Local Similarity 75.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 MLAQERR 8
DB      70 VLAAKERR 77

RESULT 7
US-10-080-170-561
; Sequence 561, Application US/10080170
; Publication No. US20030129601A1
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03495.0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 561
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
; US-10-080-170-561

Query Match          70.7%; Score 29; DB 16; Length 204;
Best Local Similarity 85.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2 LAAQERR 8
DB      166 LAEERR 172

RESULT 8
US-09-815-242-5576
; Sequence 5576, Application US/09815242
```

; Sequence 14955, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 14955  
; LENGTH: 202  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-14955

Query Match 75.6%; Score 31; DB 15; Length 202;  
Best Local Similarity 75.0%; Pred. No. 87;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAAQERV 9  
Db 132 LAAQEQI 139

RESULT 3  
US-10-156-761-8217  
; Sequence 8217, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 8217  
; LENGTH: 263  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-8217

Query Match 75.6%; Score 31; DB 15; Length 263;  
Best Local Similarity 66.7%; Pred. No. 11e+02;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAQERV 9  
Db 140 LLAARDRV 148

RESULT 4  
US-10-156-761-14852  
; Sequence 14852, Application US/10156761

; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 14852  
; LENGTH: 408  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-14852

Query Match 73.2%; Score 30; DB 15; Length 408;  
Best Local Similarity 77.8%; Pred. No. 2.8e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAAQERV 9  
Db 123 MAAEQRRV 131

RESULT 5  
US-10-117-846-16  
; Sequence 16, Application US/10117846  
; Publication No. US20020168673A1  
; GENERAL INFORMATION:  
; APPLICANT: Fuller, Margaret T  
; APPLICANT: Hales, Karen G.  
; APPLICANT: Santel, Ansgar H.  
; TITLE OF INVENTION: Mitofusins, Pzo Homologs and Functional  
; TITLE OF INVENTION: Derivatives Thereof  
; FILE REFERENCE: STAN-063CIP3  
; CURRENT APPLICATION NUMBER: US/10/117,846  
; CURRENT FILING DATE: 2002-04-05  
; PRIOR APPLICATION NUMBER: 09/413,285  
; PRIOR FILING DATE: 1999-10-06  
; PRIOR APPLICATION NUMBER: PCT/US00/27871  
; PRIOR FILING DATE: 2000-10-06  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16  
; LENGTH: 758  
; TYPE: PRT  
; ORGANISM: M. musculus  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (1)...(758)  
; OTHER INFORMATION: Xaa = Any Amino Acid  
US-10-117-846-16

Query Match 73.2%; Score 30; DB 14; Length 758;  
Best Local Similarity 75.0%; Pred. No. 5.2e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAAQERV 9  
Db 382 IAAQEQRV 389

RESULT 6  
US-09-764-891-4027



**THIS PAGE BLANK (USPTO)**

RESULT 14  
US-09-252-991A-25723  
; Sequence 25723, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 25723  
; LENGTH: 575  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-25723  
Query Match 70.7%; Score 29; DB 4; Length 575;  
Best Local Similarity 75.0%; Pred. No. 2.1e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 2 LAAQERRV 9  
|:||||:|  
DB 425 LSAQERQV 432  
RESULT 15  
US-09-082-059-2  
; Sequence 2, Application US/09082059A  
; Patent No. 6225086  
; GENERAL INFORMATION:  
; APPLICANT: Morrow, Jon S.  
; APPLICANT: Devarajan, Prasad  
; TITLE OF INVENTION: No. 6225086el Ankyrin Proteins and a Method for Their Identif  
; FILE REFERENCE: 44574-5002-US  
; CURRENT APPLICATION NUMBER: US/09/082,059A  
; CURRENT FILING DATE: 1998-05-21  
; EARLIER APPLICATION NUMBER: 60/047356  
; EARLIER FILING DATE: 1997-05-21  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 1088  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-082-059-2  
Query Match 70.7%; Score 29; DB 3; Length 1088;  
Best Local Similarity 87.5%; Pred. No. 4.1e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 LAAQERRV 9  
|:||||:|  
DB 324 LAAQEDRV 331  
Search completed: October 7, 2003, 13:34:08  
Job time : 5.89873 secs

LENGTH: 399 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-023-591A-2  
Query Match 70.7%; Score 29; DB 3; Length 399;  
Best Local Similarity 55.6%; Pred. No. 1.4e+02;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MLLAQERRV 9  
|:||||:|  
DB 382 MLLAARKKRI 390  
RESULT 12  
US-09-351-150A-5  
; Sequence 5, Application US/09351150A  
; Patent No. 6538177  
; GENERAL INFORMATION:  
; APPLICANT: Duvick, J  
; TITLE OF INVENTION: Compositions and Methods for Fumonisin Detoxification  
; FILE REFERENCE: 035718/167676, 5718-62  
; CURRENT APPLICATION NUMBER: US/09/351,150A  
; CURRENT FILING DATE: 1998-07-12  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 429  
; TYPE: PRT  
; ORGANISM: Bacterium 2412.1  
US-09-351-150A-5  
Query Match 70.7%; Score 29; DB 4; Length 429;  
Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MLLAQERR 8  
|:||||:|  
DB 1 MTAQERRR 8  
RESULT 13  
US-09-252-991A-23070  
; Sequence 23070, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 23070  
; LENGTH: 472  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-23070  
Query Match 70.7%; Score 29; DB 4; Length 472;  
Best Local Similarity 75.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MLLAQERR 8  
|:||||:|  
DB 264 LLLAQHRR 271



QY 2 LAAQERV 9  
Db 244 LAARERKV 251

RESULT 9  
PCT-US93-08528-12  
; Sequence 12, Application PC/TUS9308528  
; GENERAL INFORMATION:  
; APPLICANT: New York University  
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN  
; NUMBER OF INVENTIONS: 348  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NEIMARK  
; STREET: 419 Seventh Street, N.W., Suite 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/08528  
; FILING DATE: 09-SEP-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/943,236  
; FILING DATE: 10-SEP-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Townsend, Kevin G.  
; REGISTRATION NUMBER: 34,033  
; REFERENCE/DOCKET NUMBER: MURPHY-2 PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-628-5197  
; TELEFAX: 202-737-3528  
; TELEX: 248633  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 333 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
PCT-US93-08528-12

Query Match 70.7%; Score 29; DB 5; Length 333;  
Best Local Similarity 75.0%; Pred. No. 1.1e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAAQERV 9  
Db 244 LAARERKV 251

RESULT 10  
US-08-839-581A-2  
; Sequence 2, Application US/08839581A  
; Patent No. 5958705  
; GENERAL INFORMATION:  
; APPLICANT: Staunton, Donald E.  
; APPLICANT: Lipsky, Brian P.  
; TITLE OF INVENTION: Cytoplasmic Modulators of Integrin  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker, 6300 Sears Tower  
; CITY: Chicago  
; STATE: Illinois

; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/839,581A  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Young J. Suh  
; REGISTRATION NUMBER: P-41,337  
; REFERENCE/DOCKET NUMBER: 27866/33886  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 399 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-839-581A-2

Query Match 70.7%; Score 29; DB 2; Length 399;  
Best Local Similarity 55.6%; Pred. No. 1.4e+02;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAQERV 9  
Db 382 MAAARKRI 390

RESULT 11  
US-09-023-591A-2  
; Sequence 2, Application US/09023591A  
; Patent No. 6210914  
; GENERAL INFORMATION:  
; APPLICANT: Staunton, Donald E.  
; APPLICANT: Lipsky, Brian P.  
; TITLE OF INVENTION: Cytoplasmic Modulators of Integrin  
; TITLE OF INVENTION: Binding/Signaling  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker, 6300 Sears Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/023,591A  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/839,581  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Young J. Suh  
; REGISTRATION NUMBER: P-41,337  
; REFERENCE/DOCKET NUMBER: 27866/33886  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICANT: PatentIn Release #1.0, Version #1.30  
FILING DATE: 05-Sep-1996  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION NUMBER: US 60/074,788  
FILING DATE: 1998-02-18  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 18807  
LENGTH: 321  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-18807

Query Match 70.7%; Score 29; DB 4; Length 321;  
Best Local Similarity 75.0%; Pred. No. 1.1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAAQERR 8  
| | | | |  
Db 152 MLALQDRR 159

RESULT 8  
US-08-118-270-12  
Sequence 12, Application US/08118270  
Patent No. 5508384  
GENERAL INFORMATION:  
APPLICANT: Murphy, Randall B.  
APPLICANT: Schuster, David I.  
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN  
TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF  
NUMBER OF SEQUENCES: 348  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROWDY AND NEIMARK  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICANT: PatentIn Release #1.0, Version #1.25  
FILING DATE: 09-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/943,236  
FILING DATE: 10-SEP-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Townsend, Kevin G.  
REGISTRATION NUMBER: 34,033  
REFERENCE/DOCKET NUMBER: MURPHY-2A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 333 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-118-270-12

Query Match 70.7%; Score 29; DB 1; Length 333;  
Best Local Similarity 75.0%; Pred. No. 1.1e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICANT: PatentIn Release #1.0, Version #1.30  
FILING DATE: 05-Sep-1996  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION NUMBER: US 60/074,788  
FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 24539  
LENGTH: 272  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-24539

Query Match 70.7%; Score 29; DB 4; Length 272;  
Best Local Similarity 75.0%; Pred. No. 92;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAAQRRV 9  
| | | | |  
Db 79 LAAQRRV 86

RESULT 7  
US-09-252-991A-18807  
Sequence 18807, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICANT: PatentIn Release #1.0, Version #1.30  
FILING DATE: 05-Sep-1996  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION NUMBER: US 60/074,788  
FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 24539  
LENGTH: 272  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-24539

Query Match 70.7%; Score 29; DB 4; Length 272;  
Best Local Similarity 75.0%; Pred. No. 92;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAAQRRV 9  
| | | | |  
Db 79 LAAQRRV 86

RESULT 7  
US-09-252-991A-18807  
Sequence 18807, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.

```
Query Match      78.0%; Score 32; DB 1; Length 103;
Best Local Similarity 100.0%; Pred. No. 7.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 AAOQERV 9
Db      31 AAOQERV 37
|||||

RESULT 2
US-09-252-991A-27953
; Sequence 27953, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27953
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27953

Query Match      73.2%; Score 30; DB 4; Length 451;
Best Local Similarity 66.7%; Pred. No. 97;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 MAAQERV 9
Db      259 LVIAQERV 267
:: |||||

RESULT 3
US-09-252-991A-31942
; Sequence 31942, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31942
; LENGTH: 662
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31942

Query Match      73.2%; Score 30; DB 4; Length 662;
Best Local Similarity 75.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 MAAQERV 8
Db      222 VLAQERV 229
|||||

RESULT 4
```

```
US-08-708-573F-10
; Sequence 10, Application US/08708573F
; Patent No. 6573362
; GENERAL INFORMATION:
; APPLICANT: KOLANUS, Waldemar
; TITLE OF INVENTION: CYTOHESIN-PH PEPTIDES THAT AFFECT THE
; ABILITY OF INTEGRINS TO ADHERE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & McAuliffe
; STREET: 1666 K Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/708,573F
; FILING DATE: 05-Sep-1996
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 19534120.1
; FILING DATE: 14-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: GRANADOS, Patricia D.
; REGISTRATION NUMBER: 33,683
; REFERENCE/DOCKET NUMBER: 38005-0044
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)912-2000
; TELEFAX: (202)912-2020
; TELEX: <unknown>
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 263 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-08-708-573F-10

Query Match      70.7%; Score 29; DB 4; Length 263;
Best Local Similarity 55.6%; Pred. No. 89;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      1 MAAQERV 9
Db      249 MAAKKRI 257
|||||

RESULT 5
US-08-708-573F-14
; Sequence 14, Application US/08708573F
; Patent No. 6573362
; GENERAL INFORMATION:
; APPLICANT: KOLANUS, Waldemar
; TITLE OF INVENTION: CYTOHESIN-PH PEPTIDES THAT AFFECT THE
; ABILITY OF INTEGRINS TO ADHERE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & McAuliffe
; STREET: 1666 K Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:21:07 ; Search time 4.89873 Seconds  
(without alignments)  
77.734 Million cell updates/sec

Title: US-09-807-512-26  
Perfect score: 41  
Sequence: 1 MAAQERRV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA: \*  
1: /cgn2.6/ptodata/1/laa/5A-COMB.pep: \*  
2: /cgn2.6/ptodata/1/laa/5B-COMB.pep: \*  
3: /cgn2.6/ptodata/1/laa/6A-COMB.pep: \*  
4: /cgn2.6/ptodata/1/laa/6B-COMB.pep: \*  
5: /cgn2.6/ptodata/1/laa/PCTUS-COMB.pep: \*  
6: /cgn2.6/ptodata/1/laa/backfiles1.pep: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	32	78.0	103	1	US-08-552-142A-13
2	30	73.2	451	4	US-09-252-991A-27953
3	30	73.2	662	4	US-09-252-991A-31942
4	29	70.7	263	4	US-08-708-573F-10
5	29	70.7	288	4	US-08-708-573F-14
6	29	70.7	272	4	US-09-252-991A-24539
7	29	70.7	321	4	US-09-252-991A-18807
8	29	70.7	333	1	US-08-118-270-12
9	29	70.7	333	5	PCT-US93-08528-12
10	29	70.7	399	2	US-08-839-581A-2
11	29	70.7	399	3	US-09-023-591A-2
12	29	70.7	429	4	US-09-351-150A-5
13	29	70.7	472	4	US-09-252-991A-23070
14	29	70.7	575	4	US-09-252-991A-25723
15	29	70.7	1088	3	US-09-082-059-2
16	28	68.3	34	1	US-08-129-089-6
17	28	68.3	34	1	US-08-129-089-7
18	28	68.3	34	1	US-08-454-444-2
19	28	68.3	34	1	US-08-454-444-3
20	28	68.3	34	5	PCT-US93-01135-6
21	28	68.3	34	5	PCT-US93-01135-7
22	28	68.3	58	4	US-09-358-383C-29
23	28	68.3	150	1	US-08-362-453-10
24	28	68.3	150	1	US-08-362-453-11
25	28	68.3	150	1	US-08-362-453-12
26	28	68.3	165	1	US-08-024-330-1
27	28	68.3	165	1	US-07-952-840-1

28	28	68.3	165	1	US-08-356-021-1	Sequence 1, Appl
29	28	68.3	165	1	US-08-362-453-13	Sequence 13, Appl
30	28	68.3	165	1	US-08-362-453-14	Sequence 14, Appl
31	28	68.3	165	1	US-08-249-671A-5	Sequence 5, Appl
32	28	68.3	165	3	US-09-216-500-1	Sequence 1, Appl
33	28	68.3	165	4	US-09-730-464-1	Sequence 1, Appl
34	28	68.3	165	5	PCT-US94-01729-1	Sequence 1, Appl
35	28	68.3	166	1	US-08-362-453-9	Sequence 9, Appl
36	28	68.3	166	3	US-08-819-238A-1	Sequence 13, Appl
37	28	68.3	166	3	US-08-819-238A-13	Sequence 13, Appl
38	28	68.3	166	4	US-09-339-913B-78	Sequence 78, Appl
39	28	68.3	166	4	US-09-339-913B-80	Sequence 80, Appl
40	28	68.3	166	4	US-09-339-904A-78	Sequence 78, Appl
41	28	68.3	166	4	US-09-339-904A-80	Sequence 80, Appl
42	28	68.3	166	4	US-09-379-434-1	Sequence 1, Appl
43	28	68.3	166	4	US-09-379-434-13	Sequence 13, Appl
44	28	68.3	166	4	US-08-769-062B-78	Sequence 78, Appl
45	28	68.3	166	4	US-08-769-062B-80	Sequence 80, Appl

ALIGNMENTS

RESULT 1  
US-08-552-142A-13  
; Sequence 13, Application US/08552142A  
; Patent No. 5695995  
; GENERAL INFORMATION:  
; APPLICANT: Weintraub, Harold M.  
; APPLICANT: Lee, Jacqueline E.  
; APPLICANT: Tapscott, Stephen J.  
; APPLICANT: Hollenberg, Stanley M.  
; TITLE OF INVENTION: Neurogenic Differentiation (Neurod) Genes  
; TITLE OF INVENTION: and Proteins  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Christensen O'Connor Johnson Kindness PLLC  
; STREET: 1420 Fifth Avenue, Suite 2800  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101-2347  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/552,142A  
; FILING DATE: 02-NOV-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/239,238  
; FILING DATE: 06-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/US95/05741  
; FILING DATE: 08-MAY-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Broderick, Thomas F.  
; REGISTRATION NUMBER: 31,332  
; REFERENCE/DOCKET NUMBER: FHCR-1-8933  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 206-682-8100  
; TELEFAX: 206-225-0709  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 103 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; US-08-552-142A-13

**THIS PAGE BLANK (USPTO)**

```

RL Advances in chemical conversions for mitigating carbon dioxide,
RL pp.114:653-656, Elsevier Science, Amsterdam (1998).
DR EMBL; AB015804; BAA77599.1; -.
DR Hypothetical protein.
KW SEQUENCE 215 AA; 23925 MW; ABC329BB939205F1 CRC64;

Query Match 75.6%; Score 31; DB 2; Length 215;
Best Local Similarity 66.7%; Pred. No. 94;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAAQERRV 9
:|||||
Db 36 LLAARERRI 44

RESULT 14
Q9ALR3 PRELIMINARY; PRT; 368 AA.
AC Q9ALR3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE CztB.
GN CztB.
OS Pseudomonas fluorescens.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=294;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13525;
RA Feng S.F., Rosbach S.;
RT "A locus involved in metal homeostasis in Pseudomonas fluorescens
RT encodes a proton/cation antiporter of the RND family and a two-
RT component system."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY007258; AAC09628.1; -.
DR InterPro; IPR006143; HlyD.
DR Pfam; PF00529; HlyD; 1.
SQ SEQUENCE 368 AA; 39629 MW; 99B35B74F691158F CRC64;

Query Match 75.6%; Score 31; DB 2; Length 368;
Best Local Similarity 87.5%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAAQERRV 9
||| ||||
Db 112 LAASERV 119

RESULT 15
P72346 PRELIMINARY; PRT; 495 AA.
AC P72346;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE L-amino acid oxidase.
GN AOX.
OS Synecococcus sp. (strain PCC 6301) (Anacystis nidulans).
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
OX NCBI_TaxID=1139;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96138550; PubMed=8547316;
RA Bockholt R., Masepohl B., Kruft V., Wittmann-Liebold B., Pistorius E.;
RT "Partial amino acid sequence of an L-amino acid oxidase from the
RT cyanobacterium Synecococcus PCC6301, cloning and DNA sequence
RT analysis of the aoxA gene."
RL Biochim. Biophys. Acta 1264:289-293(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Heindl A., Gau A.E., Vallon O., Kruft V., Wittmann-Liebold B.,
```

```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue-Lung, and Uterus;
RA Strausberg R.;
RL EMBL; BC017235; AHI17235.1; -
DR EMBL; BC014918; AHI14918.1; -
KW Hypothetical protein.
SQ SEQUENCE 631 AA; 70737 MW; 4EFC3AF5D5BBABD3 CRC64;

Query Match 78.0%; Score 32; DB 4; Length 631;
Best Local Similarity 87.5%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAAQERR 8
Db 264 MLAAQSRR 271
|||||

RESULT 11
Q8LIY9
ID Q8LIY9 PRELIMINARY; PRT; 829 AA.
AC Q8LIY9
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE OSJNBa0014K08.24 protein.
GN OSJNBa0014K08.24
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoideae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC
clone:OSJNBa0014K08.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003376; BAC05597.1; -
DR Gramineae; Q8LIY9; -
DR InterPro; IPR001993; Mitoch_carrier.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF04094; DUF390; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
SQ SEQUENCE 829 AA; 90728 MW; DC8F6B438C1EF1FC CRC64;

Query Match 78.0%; Score 32; DB 10; Length 829;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 AAQERRV 9
Db 584 AAQERRV 590
|||||

RESULT 12
Q9RD77
ID Q9RD77 PRELIMINARY; PRT; 205 AA.
AC Q9RD77
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative TetR-family transcriptional regulatory protein.
GN SCO0800 OR SCF43.11C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Saunders D.C., Harris D.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Redenbach M., Kieser H.M., Denapante D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
RL Nature 417:141-147(2002).
CC -1- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
DR EMBL; AL539106; CAB66200.1; -
DR InterPro; IPR001647; HTH_Tetr.
DR Pfam; PF00440; tetr; 1.
DR PRINTS; PR00455; HTH_TETR.
DR DNA-binding; Transcription regulation; Complete proteome.
KW DNA-binding; Transcription regulation; Complete proteome.
SQ SEQUENCE 205 AA; 23066 MW; E8B82329FF445F1 CRC64;

Query Match 75.6%; Score 31; DB 16; Length 205;
Best Local Similarity 75.0%; Pred. No. 89;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAAQERRV 9
Db 135 LAAQEROI 142
|||||

RESULT 13
Q9WX74
ID Q9WX74 PRELIMINARY; PRT; 215 AA.
AC Q9WX74
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical 23.9 kDa protein (ORF216 protein).
GN ORF216.
OS Acetobacter xylinus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC Acetobacteraceae; Gluconacetobacter.
OX NCBI_TaxID=28448;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM7664;
RA Umeda Y., Hirano A., Hon-nami K., Kunito S., Akiyama H., Onizuka T.,
RA Ikeuchi M., Inoue Y.;
RT "Conversion of CO2 into cellulose by gene manipulation of microalgae:
cloning of cellulose synthase genes from acetobacter xylinum.";
RL (In) Inui T., Anpo M., Izui K., Yanagida S., Yamaguchi T. (eds.);

```

```

Best Local Similarity 100.0%; Pred. No. 55;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAQERR 9
Db 53 AAQERR 59

RESULT 6
O14710 PRELIMINARY; PRT; 360 AA.
AC O14710;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Cell cycle progression 2 protein.
GN CPR2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98043401; PubMed=9383053;
RA Edwards M.C., Liegeois N., Horecka J., Depinho R.A., Sprague G.F. Jr.,
RA Tyers M., Elledge S.J.;
RT "Human CYP (cell cycle progression restoration) genes impart a Far-
RT phenotype on yeast cells.";
RL Genetics 147:1063-1076(1997).
DR EMBL; AF011792; AAB69312.1; -.
SQ SEQUENCE 360 AA; 40254 MW; 9ACOF8E46DA11199 CRC64;

Query Match 78.0%; Score 32; DB 4; Length 360;
Best Local Similarity 87.5%; Pred. No. 95;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAAQERR 8
Db 37 MLAAQERR 44

RESULT 7
O05989 PRELIMINARY; PRT; 413 AA.
AC O05989;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Lysostaphin immunity factor.
GN LIF.
OS Staphylococcus simulans.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1286;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97260121; PubMed=9106216;
RA Thumm G., Gotz F.;
RT "Studies on polysostaphin processing and characterization of the
RT lysostaphin immunity factor (lif) of Staphylococcus simulans biovar
RT staphylolyticus.";
RL Mol. Microbiol. 23:1251-1265(1997).
DR EMBL; U66883; AAB53784.1; -.
DR InterPro; IPR003447; Meth.resist.
DR Pfam; PF02388; FemAB; 1.
SQ SEQUENCE 413 AA; 49121 MW; B183532C6D96F551 CRC64;

Query Match 78.0%; Score 32; DB 2; Length 413;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAAQERR 9
Db 289 LAAQERR 296

us-09-807-512-26.rspt

RESULT 8
Q9BUC6 PRELIMINARY; PRT; 456 AA.
AC Q9BUC6;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Similar to cell cycle progression 2 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Uterus;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC002732; AAH02732.1; -.
SQ SEQUENCE 456 AA; 51368 MW; 19D4BIAC5A47BA75 CRC64;

Query Match 78.0%; Score 32; DB 4; Length 456;
Best Local Similarity 87.5%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAAQERR 8
Db 89 MLAAQERR 96

RESULT 9
Q8NDM4 PRELIMINARY; PRT; 631 AA.
AC Q8NDM4;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Hypothetical protein.
GN DKF2P727A151.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Breast cancer;
RA Poustka A., Klein M., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL833840; CAD38700.1; -.
KW Hypothetical protein.
SQ SEQUENCE 631 AA; 70770 MW; 1199508B6CB8A03 CRC64;

Query Match 78.0%; Score 32; DB 4; Length 631;
Best Local Similarity 87.5%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAAQERR 8
Db 264 MLAAQERR 271

RESULT 10
Q96920 PRELIMINARY; PRT; 631 AA.
AC Q96920;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```



QY 1 MAAQERRV 9  
 Db 17 MAAQERRV 25

## RESULT 2

Q95987 ID Q95987 PRELIMINARY; PRT; 109 AA.  
 AC Q95987;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE CML-recognized antigen ON melanoma (CAMEL).  
 GN CAMEL.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Melanoma;  
 RA Aarnoudse C.A., Doel vanden P.B., Heemskerk B., Schrier P.I.;  
 RT "IL-2 induced melanoma-specific CTL recognize CAMEL, an unexpected  
 translation product of LAGE-1";  
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ012835; CAA10197.1; -;  
 DR EMBL; AJ012833; CAA10193.1; -;  
 DR EMBL; AJ012834; CAA10195.1; -;  
 SQ SEQUENCE 109 AA; 11689 MW; 188F85BC04C1F5F0 CRC64;

Query Match 100.0%; Score 41; DB 4; Length 109;  
 Best Local Similarity 100.0%; Pred. No. 0.32;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAQERRV 9  
 Db 17 MAAQERRV 25

## RESULT 3

Q8X0T0 ID Q8X0T0 PRELIMINARY; PRT; 655 AA.  
 AC Q8X0T0;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Hypothetical 71.6 kDa protein.  
 GN 18f11.080.  
 OS Neurospora crassa.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariales; Sordariaceae; Neurospora.  
 OX NCBI\_TaxID=5141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,  
 RA Nyakatura G., Mewes H.W., Mannhaupt G.;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA German Neurospora genome project;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL670011; CAD21413.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 655 AA; 71619 MW; 159ADE04A519425F CRC64;

Query Match 82.9%; Score 34; DB 3; Length 655;  
 Best Local Similarity 87.5%; Pred. No. 63;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAAQERRV 9  
 Db 265 IAAQERRV 272

## RESULT 4

Q8H8L8 ID Q8H8L8 PRELIMINARY; PRT; 129 AA.  
 AC Q8H8L8;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Hypothetical protein.  
 GN OSJNBA0070N04.25.  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,  
 RA Overton II L.L., Tsitrin T., Kim M.M., Bera J.J., Jin S.S.,  
 RA Fanrosh D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,  
 RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,  
 RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,  
 RA White O., Salzberg S.L., Fraser C.M.;  
 RT "Oryza sativa chromosome 3 BAC OSJNBA0070N04 genomic sequence.";  
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC091494; AAN65022.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 129 AA; 13913 MW; 073ED4F410DE5D51 CRC64;

Query Match 80.5%; Score 33; DB 10; Length 129;  
 Best Local Similarity 87.5%; Pred. No. 21;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAQERR 8  
 Db 84 VLAQERR 91

## RESULT 5

Q9L1Z6 ID Q9L1Z6 PRELIMINARY; PRT; 208 AA.  
 AC Q9L1Z6;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Putative RNA polymerase sigma factor.  
 GN SCO2639 OR SC8E4A.09C.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2) / M145;  
 RX MEDLINE=21996410; PubMed=12000953;  
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,  
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 coelicolor A3(2).";  
 RT Nature 417:141-147(2002).  
 DR EMBL; AL939113; CAB71814.1; -;  
 KW Complete proteome.  
 SQ SEQUENCE 208 AA; 22978 MW; A535200CE7DC603D CRC64;

Query Match 78.0%; Score 32; DB 16; Length 208;

17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45

## RESULT, T 1

REPORT: 58 AA.

O95146:					Created)
AC	TREMBLrel.	10,			(TREMBLrel.
DT	01-MAY-1999				01-MAY-1999
DT	01-MAY-1999				01-MAY-1999
DT	01-MAR-2001				01-MAR-2001
DE	LAGE-2ALT protein isoform.				LAGE-2ALT protein isoform.

GN  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
LAGE-2.  
GN

ppred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and  $\sigma^2$  is the variance of the total score distribution.

## SUMMARY

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	41	100.0	58	4	Q95146	homo sapien
2	41	100.0	109	4	Q95987	homo sapien
3	34	82.9	655	3	Q8X0T0	neurospora
4	33	80.5	129	10	Q8H8L8	oryza sativ
5	32	78.0	208	16	Q9L1Z6	streptomyce
6	32	78.0	360	4	Q14710	homo sapien
7	32	78.0	413	2	Q95989	staphylococ
8	32	78.0	456	4	Q9BUC6	homo sapien
9	32	78.0	631	4	Q8NDM4	homo sapien
10	32	78.0	631	4	Q96320	homo sapien
11	32	78.0	829	10	Q8LIY9	oryza sativ
12	31	75.6	205	16	Q9RD77	streptomyce
13	31	75.6	215	2	Q9WK74	acetobacter
14	31	75.6	368	2	Q9ALR3	psuedomonas
15	31	75.6	495	2	P72346	synechococc
16	31	75.6	500	16	Q9A3K8	caulobacter

Best Local Similarity 75.0%; Pred. No. 77;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAAQERRV 9  
      |||||  
Db 451 LAARERRL 458

## RESULT 15

YK03\_YEAST  
ID YK03\_YEAST STANDARD; PRT; 530 AA.  
AC P36119;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE Hypothetical 60.8 kDa protein in YPT52-DBP7 intergenic region.  
GN YKR023W.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288c;  
RA Duesterhoeft A., Moestl D., Poehlmann R., Philipssen P.;  
RL Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; Z28248; CAA82095.1; -  
DR PIR; S38092; S38092  
DR SGD; S0001731; YKR023W.  
KW Hypothetical protein.  
SQ SEQUENCE 530 AA; 60799 MW; 57515FA658E7535A CRC64;

Query Match 70.7%; Score 29; DB 1; Length 530;  
Best Local Similarity 85.7%; Pred. No. 83;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAAQER 7  
      |||||  
Db 325 LLAAQER 331

Search completed: October 7, 2003, 13:26:28  
Job time : 3.62025 secs

70 7% Score 29: DB 1: Length 494;

RT domains";  
 RL mature 384:481-484(1996).  
 [2]  
 RN SEQUENCE FROM N.A. (ISOFORM 2), AND CHARACTERIZATION.  
 RP TISSUE=Brain;  
 RC MEDLINE=98079021; PubMed=9417041;  
 RA Frank S.F., Upender S.K., Hansen S.H., Casanova J.E.;  
 RT "ARNO is a guanine nucleotide exchange factor for ADP-ribosylation  
 factor 6";  
 RL J. Biol. Chem. 273:23-27(1998).  
 [3]  
 RN SEQUENCE FROM N.A. (ISOFORM 1).  
 RP TISSUE=Lung;  
 RC MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Pahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 [4]  
 RN X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 51-252.  
 RP MEDLINE=98135767; PubMed=9476900;  
 RA Mossessova E., Gulbis J.M., Goldberg J.;  
 RT "Structure of the guanine nucleotide exchange factor Sec7 domain of  
 human ARNO and analysis of the interaction with ARF GTPase.";  
 RL Cell 92:415-423(1998).  
 [5]  
 RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 51-252.  
 RP MEDLINE=98169075; PubMed=9510256;  
 RA Cherfils J., Menetrey J., Mathieu M., le Bras G., Robineau S.,  
 RA Beraud-Dufour S., Antony B., Chardin P.;  
 RT "Structure of the Sec7 domain of the Arf exchange factor ARNO.";  
 RL Nature 392:101-105(1998).  
 CC -!- FUNCTION: PROMOTES GUANINE-NUCLEOTIDE EXCHANGE ON ARF1, ARF3 AND  
 CC ARF6. PROMOTES THE ACTIVATION OF ARF THROUGH REPLACEMENT OF GDP  
 CC WITH GTP.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=Q99418-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q99418-2; Sequence=VSP\_006036;  
 CC -!- TISSUE SPECIFICITY: Ubiquitous.  
 CC -!- SIMILARITY: Contains 1 SEC7 domain.  
 CC -!- SIMILARITY: Contains 1 PH domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X99753; CAA68084.1; -;  
 CC EMBL; U70728; AAB09591.1; -;  
 CC EMBL; BC004361; AAH04361.1; -;  
 CC PDB; 1PBV; 09-MAR-99.

DR Genew: HGNC:9502; PSCD2.  
 DR MIM; 602488; -;  
 DR GO; GO:0005624; C:membrane fraction; TAS.  
 DR GO; GO:0005886; C:plasma membrane; TAS.  
 DR GO; GO:0005086; F:ARF guanyl-nucleotide exchange factor activity; TAS.  
 DR GO; GO:0007012; P:actin cytoskeleton reorganization; TAS.  
 DR GO; GO:0006897; P:apoptosis; TAS.  
 DR InterPro; IPR001849; PH.  
 DR InterPro; IPR000904; Sec7.  
 DR Pfam; PF00169; PH; 1.  
 DR Pfam; PF01369; Sec7; 1.  
 DR SMART; SM00233; PH; 1.  
 DR SMART; SM00222; Sec7; 1.  
 DR PROSITE; PS00003; PH\_DOMAIN; 1.  
 DR PROSITE; PS00190; SEC7; 1.  
 KW Guanine-nucleotide releasing factor; Coiled coil;  
 KW Alternative splicing; 3D-structure.  
 FT DOMAIN 10 63 COILED COIL (POTENTIAL).  
 FT DOMAIN 72 201 SEC7.  
 FT DOMAIN 259 376 PH.  
 FT VARSPLIC 272 272 Missing (in isoform 2).  
 FT /FTID=VSP\_006036.  
 FT HELIX 54 74  
 FT HELIX 76 85  
 FT TURN 86 87  
 FT HELIX 93 102  
 FT TURN 104 105  
 FT HELIX 108 115  
 FT TURN 116 116  
 FT HELIX 120 131  
 FT TURN 132 132  
 FT TURN 136 137  
 FT HELIX 140 149  
 FT HELIX 157 174  
 FT TURN 176 178  
 FT HELIX 182 200  
 FT TURN 202 203  
 FT HELIX 210 216  
 FT TURN 217 220  
 FT STRAND 221 221  
 FT TURN 222 223  
 FT STRAND 224 224  
 FT HELIX 228 240  
 SQ SEQUENCE 400 AA; 46546 MW; 70441A58483BD0E1 CRC64;  
 Query Match 70.7%; Score 29; DB 1; Length 400;  
 Best Local Similarity 55.6%; Pred. No. 62;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MLAAQERRY 9  
 Db 383 MLAARKKRI 391  
 RESULT 13  
 CYH2\_MOUSE  
 ID CYH2\_MOUSE STANDARD; PRT; 400 AA.  
 AC P97695; O89099;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Cytohesin 2 (ARF nucleotide-binding site opener) (ARNO protein) (CLM2)  
 DE (SEC7 homolog B) (msec7-2).  
 GN PSCD2 OR SEC7B.  
 OS Mus musculus (Mouse), and  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090, 10116;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).  
 RC SPECIES=Mouse; TISSUE=Brain;  
 RX MEDLINE=98416066; PubMed=9744817;

FT CONFLICT 85 85 G -> R (IN REF. 1).  
FT CONFLICT 107 107 E -> G (IN REF. 1).  
FT CONFLICT 194 194 R -> S (IN REF. 1).  
FT CONFLICT 237 237 G -> A (IN REF. 1).  
SQ SEQUENCE 267 AA; 28867 MW; 49D64AD5A5414AFD CRC64;

Query Match 70.7%; Score 29; DB 1; Length 267;

Best Local Similarity 85.7%; Pred. No. 42;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAQERRV 9  
||:||||  
DB 34 AAERRV 40

## RESULT 10

HCHA\_STAAM STANDARD; PRT; 292 AA.

AC Q99W58;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DE 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Chaperone protein hcha (Hsp31).  
GN HCHA OR SAV0351 OR SA0309.  
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and  
OS Staphylococcus aureus (strain N315).  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
OX NCBI\_TaxID=158878, 158879;  
[1]

## SEQUENCE FROM N.A.

RC STRAIN=Mu50 / ATCC 700699, and N315;  
RX MEDLINE=213111952; PubMed=11418146;  
RA Kuroda M., Ohca T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,  
Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,  
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,  
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,  
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,  
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,  
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;  
RT "Whole genome sequencing of methicillin-resistant Staphylococcus  
aureus."  
RL Lancet 357:1225-1240(2001).

-!- FUNCTION: Seems to use temperature-induced exposure of structured  
hydrophobic domains to capture early unfolding protein  
intermediates and rapidly release them in an active form once  
stress has abated (by similarity).

-!- SIMILARITY: Belongs to the hcha family.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
EMBL: AP003359; BAB56713.1; -;  
DR EMBL: AP003130; BAB41740.1; -;  
DR PIR: A89823; A89823.

DR HAMAP: MF\_01046; -; 1.

DR InterPro: IPR002818; ThIJ.

DR Pfam: PF01965; DJ-1.Pfpl; 1.

KW Chaperone; Complete proteome.

SQ SEQUENCE 292 AA; 32176 MW; 7ADA2635287373CD CRC64;

Query Match 70.7%; Score 29; DB 1; Length 292;

Best Local Similarity 66.7%; Pred. No. 46;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAAQERRV 9  
|:|:|:|:|  
DB 57 MIAAERYV 65

## RESULT 11

HCHA\_STAAM STANDARD; PRT; 292 AA.

AC Q8NXY2;  
DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Chaperone protein hcha (Hsp31).

GN HCHA OR MW0506.

OS Staphylococcus aureus (strain MW2).

OC Bacteria; Firmicutes; Bacillales; Staphylococcus.

OX NCBI\_TaxID=196620;

RN [1]

SEQUENCE FROM N.A.

RX MEDLINE=22040717; PubMed=12044378;

RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,

RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,

RA Yamamoto K., Hiramatsu K.;

RT "Genome and virulence determinants of high virulence community-

acquired MRSA."

RL Lancet 359:1819-1827(2002).

CC -!- FUNCTION: Seems to use temperature-induced exposure of structured

hydrophobic domains to capture early unfolding protein

intermediates and rapidly release them in an active form once

stress has abated (by similarity).

CC -!- SIMILARITY: Belongs to the hcha family.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
EMBL: AP004823; BAB94371.1; -;

DR HAMAP: MF\_01046; -; 1.

DR InterPro: IPR002818; ThIJ.

DR Pfam: PF01965; DJ-1.Pfpl; 1.

KW Chaperone; Complete proteome.

SQ SEQUENCE 292 AA; 32180 MW; CA9E227128741960 CRC64;

## Query Match

Best Local Similarity 70.7%; Score 29; DB 1; Length 292;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAAQERRV 9  
|:|:|:|:|  
DB 57 MIAAERYV 65

## RESULT 12

CHH2\_HUMAN

ID CYH2\_HUMAN STANDARD; PRT; 400 AA.

AC Q99418; Q92958;

DT 15-JUL-1998 (Rel. 36, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Cytohesin 2 (ARF nucleotide-binding site opener) (ARNO protein) (ARF

exchange factor).

GN PSCD2 OR ARNO.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

SEQUENCE FROM N.A. (ISOFORM 1).

RP TISSUE=Brain;

RX MEDLINE=97100951; PubMed=8945478;

RA Chardin P., Paris S., Antony B., Robineau S., Bernaud-Dufour S.,

RA Jackson C.L., Chabre M.;

RT "A human exchange factor for ARF contains Sec7-and pleckstrin-homology

RN SEQUENCE FROM N.A.  
 RX MEDLINE=93326628; PubMed=8101453;  
 RA Hodgkinson J.E., Davidson C.L., Beresford J., Sharpe P.T.;  
 RT "Expression of a human homeobox-containing gene is regulated by  
 RL 1,25(OH)2D3 in bone cells.";  
 RN Blochim. Biophys. Acta 1174:11-16(1993).  
 [2]  
 RN SEQUENCE FROM N.A., AND VARIANT CRS2 HIS-148.  
 RX MEDLINE=94037092; PubMed=8106171;  
 RA Jabs E.W., Ma L., Li X., Mueller U., Sparkes R.S., Luo W.,  
 RA Jackson C.E., Warman M.L., Mulliken J.B., Snead M., Haworth I.,  
 RA Maxson R.E.;  
 RT "A mutation in the homeodomain of the human MSX2 gene in a family  
 RT affected with autosomal dominant craniosynostosis.";  
 RL Cell 75:443-450(1993).  
 [3]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=95204995; PubMed=7897272;  
 RA Iimura T.;  
 RT "Molecular cloning and expression of homeobox-containing genes during  
 RT hard tissue development.";  
 RL Kokubyo Gakkai Zasshi 61:590-604(1994).  
 [4]  
 RN SEQUENCE FROM N.A.  
 RX TISSUE=Tooth;  
 MEDLINE=98182594; PubMed=9522127;  
 RA Iimura T., Takeda K., Goseki M., Maruoka Y., Sasaki S., Oida S.;  
 RT "Characterization of two length cDNA for human MSX-2 from dental  
 RT pulp-derived cells.";  
 RL DNA Seq. 8:87-92(1997).  
 [5]  
 RN SEQUENCE FROM N.A.  
 RX TISSUE=Pancreas;  
 MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins S.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong F.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Feing J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 [6]  
 RN SEQUENCE OF 96-267 FROM N.A.  
 RX MEDLINE=93326115; PubMed=7687426;  
 RA Suzuki M., Tanaka M., Iwase T., Naito Y., Sugimura H., Kinō I.;  
 RT "Over-expression of HOX-8, the human homologue of the mouse Hox-8  
 RT homeobox gene, in human tumors.";  
 RL Biochem. Biophys. Res. Commun. 194:187-193(1993).  
 [7]  
 RN VARIANTS PFM1 PRO-154 AND HIS-172.  
 RX MEDLINE=20231554; PubMed=10767351;  
 RA Wuyts W., Reardon W., Preis S., Homfray T., Rasore-Quartino A.,  
 RA Christians H., Willems P.J., Van Hul W.;  
 RT "Identification of mutations in the MSX2 homeobox gene in families  
 RT affected with foramina parietalia permagna.";  
 RL Hum. Mol. Genet. 9:1251-1255(2000).  
 [8]  
 RN VARIANTS PFM1 159-ARG-LYS-160 DEL AND HIS-172, AND VARIANT MET-129.  
 RX MEDLINE=20206562; PubMed=10742103;

RA Walkie A.O.M., Tang Z., Elanko N., Walsh S., Twigg S.R.F., Hurst J.A.,  
 RA Wall S.A., Chrzanoska K.H., Maxson R.E. Jr.;  
 RT "Functional haploinsufficiency of the human homeobox gene MSX2 causes  
 RT defects in skull ossification.";  
 RL Nat. Genet. 24:387-390(2000).  
 CC -1- FUNCTION: PROBABLE MORPHOGENETIC ROLE. MAY PLAY A ROLE IN LIMB-  
 CC PATTERN FORMATION.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- DISEASE: Defects in MSX2 are the cause of parietal foramina 1  
 CC (PFM1) [MIM:168500], also known as foramina parietalia permagna  
 CC (FPP). PFM1 is an autosomal dominant disease characterized by oval  
 CC defects of the parietal bones caused by deficient ossification  
 CC around the parietal notch, which is normally obliterated during  
 CC the fifth fetal month.  
 CC -1- DISEASE: Defects in MSX2 are the cause of craniosynostosis type 2  
 CC (CRS2) [MIM:604757], also known as craniosynostosis Boston-type  
 CC (CSB). CRS2 is characterized by the premature fusion of calvarial  
 CC sutures. This developmental anomaly causes abnormal skull shape.  
 CC -1- SIMILARITY: BELONGS TO THE MSH HOMEBOX FAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X69295; CAAG49156.1; -;  
 DR EMBL; L22499; AAB42178.1; -;  
 DR EMBL; L22498; AAB42178.1; JOINED.  
 DR EMBL; S75361; AAB33867.1; -;  
 DR EMBL; S75308; AAD14169.1; -;  
 DR EMBL; D89377; BAA13949.1; -;  
 DR EMBL; D31771; BAA06549.1; -;  
 DR EMBL; BC015509; AAL15509.1; -;  
 DR EMBL; D14970; BAA03611.1; -;  
 DR PIR; A49068; A49068.  
 DR HSP; P22808; INK3.  
 DR TRANSFAC; T02075; -;  
 DR Genew; HGNC:7392; MSX2.  
 DR MIM; 123101; -;  
 DR MIM; 168500; -;  
 DR MIM; 604757; -;  
 DR GO; GO:0001501; P:skeletal development; TAS.  
 DR InterPro; IPR001356; Homeobox.  
 DR InterPro; IPR000047; HTH\_lambdarepressr.  
 DR Pfam; PF000046; homeobox; 1.  
 DR PRINTS; PR00024; HOMEBOX.  
 DR PRINTS; PR00031; HTHREPRESSR.  
 DR ProDom; PD000010; Homeobox; 1.  
 DR SMART; SM00389; HOX; 1.  
 DR PROSITE; PS00027; HOMEBOX 1; 1.  
 DR PROSITE; PS00071; HOMEBOX 2; 1.  
 DR Homeobox; DNA-binding; Developmental protein; Nuclear protein;  
 DR Multigene family; Polymorphism; Disease mutation.  
 DR DNA\_BIND 142 201  
 DR VARIANT 129 129 T -> M.  
 DR VARIANT 148 148 P -> H (in CRS2; gain of function).  
 DR VARIANT 154 154 /FTid=VAR\_003755.  
 DR VARIANT 159 160 /FTid=VAR\_010786.  
 DR VARIANT 172 172 Missing (in PFM1; loss of function).  
 DR VARIANT 172 172 R -> H (in PFM1; loss of function).  
 DR CONFLICT 28 P -> L (IN REF. 1).  
 DR CONFLICT 32 E -> A (IN REF. 1).  
 DR CONFLICT 62 ASPL -> SPAP (IN REF. 1).  
 DR CONFLICT 67 A -> P (IN REF. 1).  
 DR CONFLICT 69 S -> G (IN REF. 1).  
 DR CONFLICT 75 T -> H (IN REF. 1).

Tue Oct 7 15:52:50 2003

us-09-807-512-26.rsp

DR SMART: SM00116; CBS: 1.  
 DR TIGRFAMS: TIGR01186; PROV: 1.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER\_1; 1.  
 DR PROSITE: PS00893; ABC\_TRANSPORTER\_2; 1.  
 KW Inner membrane; Amino-acid transporter; Transport; ATP-binding; Repeat;  
 KW CBS domain; Complete proteome.  
 FT DOMAIN 29 265 ABC\_TRANSPORTER.  
 FT DOMAIN 279 334 CBS 1.  
 FT DOMAIN 340 393 CBS 2.  
 FT NP\_BIND 61 68 ATP (BY SIMILARITY).  
 FT CONFLICT 158 158 A -> R (IN REF. 1).  
 SQ SEQUENCE 400 AA; 44124 MW; 1FDF58DE0A9A792C2 CRC64;

Query Match 73.2%; Score 30; DB 1; Length 400;  
 Best Local Similarity 85.7%; Pred. No. 37;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAAQRR 8  
 DB 136 TAAQRR 142

## RESULT 7

ID YF98\_METTH STANDARD; PRT; 140 AA.  
 AC O27635;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Hypothetical protein MTH1598.  
 GN MTH1598.  
 OS Methanobacterium thermoautotrophicum.  
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;  
 OC Methanobacteriaceae; Methanothermobacter.  
 OX NCBI\_TaxID=187420;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=Delta H;  
 RX MEDLINE=98037514; PubMed=9371463;  
 RA Smith D.R., Doucette-Stamm L.A., DeLoughery C., Lee H.-M., Dubois J.,  
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,  
 RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,  
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,  
 RA Jlwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,  
 RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,  
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;  
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
 deltaH: functional analysis and comparative genomics.";  
 RL J. Bacteriol. 179:7135-7155(1997).  
 CC -1- SIMILARITY: BELONGS TO THE UPF0211 FAMILY.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL: AE000919; AAB86071.1; -  
 PIR: C69080; C69080.  
 PDB: 1JW3; 27-FEB-02.  
 HAMAP: MF\_01222; -; 1.  
 InterPro: IPR002804; DUF101.  
 Pfam: PF01951; DUF101; 1.  
 ProDom: PD012969; DUF101; 1.  
 KW Hypothetical protein; Complete proteome; 3D-structure.  
 SQ SEQUENCE 140 AA; 16183 MW; 1F82F5705655AFC CRC64;

Query Match 70.7%; Score 29; DB 1; Length 140;  
 Best Local Similarity 85.7%; Pred. No. 22;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAOQRR 9  
 DB 45 AAOQRR 51

## RESULT 8

CYSH\_VIBVU STANDARD; PRT; 258 AA.  
 ID Q8CWR6;  
 AC Q8CWR6;  
 DT 15-SEP-2003 (Rel. 42, Created)  
 DT 15-SEP-2003 (Rel. 42, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Phosphadenosine phosphosulfate reductase (EC 1.8.4.8) (PAPS  
 reductase, thioredoxin dependent) (PADOs reductase) (3'-  
 phosphoadenylylsulfate reductase) (PAPS sulfotransferase).  
 DE Phosphadenylylsulfate reductase.  
 GN CYSH OR VV11404.  
 OS Vibrio vulnificus.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
 OC Vibrionaceae; Vibrio.  
 OX NCBI\_TaxID=672;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=CMCP6;  
 RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,  
 RA Choy H.E.;  
 RT "Complete genome sequence of Vibrio vulnificus CMCP6.";  
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Reduction of activated sulfate into sulfite.  
 CC -1- CATALYTIC ACTIVITY: Adenosine 3',5'-bisphosphate + sulfite +  
 CC oxidized thioredoxin = 3'-phosphoadenylyl sulfate + reduced  
 CC thioredoxin.  
 CC -1- PATHWAY: Sulfate activation; cysteine biosynthesis reductive  
 CC branch; third step.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity). CYSH SUBFAMILY.  
 CC -1- SIMILARITY: BELONGS TO THE PAPS REDUCTASE FAMILY.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL: AE016801; AAO09853.1; -  
 DR HAMAP: MF\_00063; -; 1.  
 DR Pfam: PF01507; PAPS\_reduct; 1.  
 DR TIGRFAMS: TIGR00434; cysh; 1.  
 KW Oxidoreductase; Cysteine biosynthesis; Complete proteome.  
 SQ SEQUENCE 258 AA; 29448 MW; C5078F647D681687 CRC64;

Query Match 70.7%; Score 29; DB 1; Length 258;  
 Best Local Similarity 85.7%; Pred. No. 40;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAAQRR 7  
 DB 35 MLAAQRR 41

## RESULT 9

MSX2\_HUMAN STANDARD; PRT; 267 AA.  
 ID MSX2\_HUMAN  
 AC P35548;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Homeobox protein MSX-2 (Hox-8).  
 GN MSX2 OR HOXB8.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;



```

DR TIGRFS: TIGR00978; asd_EA; 1.
DR PROSITE: PS01103; ASD; 1.
KW Oxidoreductase; NADP; Methionine biosynthesis; Threonine biosynthesis;
KW Amino-acid biosynthesis; Complete proteome.
FT ACT_SITE 148 BY SIMILARITY.
SQ SEQUENCE 343 AA; 37696 MW; A6C9F5D045CBFC21 CRC64;

Query Match 75.6%; Score 31; DB 1; Length 343;
Best Local Similarity 87.5%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAAQERRV 9
Db 33 LAASERV 40

RESULT 5
Y582_METMA STANDARD; PRT; 376 AA.
AC QBP2B2;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DE Hypothetical glycosyl transferase MM0582 (EC 2.4.1.10).
GN MM0582.
OS Methanosarcina mazei (Methanosarcina frisia).
OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2209;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Goel / G01 / ATCC BAA-199 / DSM 3647 / OCM 88;
RX MEDLINE=22120827; PubMed=12125824;
RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA Martinez-Arias R., Henne A., Wierer A., Baeumer S., Jacobi C.,
RA Brueggemann H., Lienard T., Christmann A., Boemcke M., Steckel S.,
RA Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA Fritz H.-J., Gottschalk G.;
RT "The genome of Methanosarcina mazei: evidence for lateral gene
RT transfer between Bacteria and Archaea."
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
CC -1- SIMILARITY: Belongs to the glycosyltransferase family 28.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB013283; AAM30278.1; -.
CC InterPro; IPR007235; Glyco_tran_28.C.
CC Pfam; PF04101; Glyco_tran_28.C; 1.
CC Hypothetical protein; Transferase; Glycosyltransferase;
KW Complete proteome.
SQ SEQUENCE 376 AA; 42120 MW; D2EAD27190BEA711 CRC64;

Query Match 73.2%; Score 30; DB 1; Length 376;
Best Local Similarity 66.7%; Pred. No. 35;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIAAQRRV 9
Db 109 ILAAQRRV 117

RESULT 6
PROV_SALTY STANDARD; PRT; 400 AA.
AC P17328;
DT 01-AUG-1990 (Rel. 15, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)

```

```

DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glycine betaine/L-proline transport ATP-binding protein prov.
GN PROV OR STM2809.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RX MEDLINE=90113884; PubMed=2691838;
RA Stirling D.A., Hulton C.S.J., Maddell L., Park S.F., Stewart G.S.A.B.,
RA Booth I.R., Higgins C.F.;
RT "Molecular characterization of the proU loci of Salmonella
RT typhimurium and Escherichia coli encoding osmoregulated glycine
RT betaine transport systems."
RL Mol. Microbiol. 3:1025-1038(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porvollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2."
RL Nature 413:852-856(2001).
RN [3]
RP SEQUENCE OF 1-291 FROM N.A.
RX MEDLINE=89359099; PubMed=2548994;
RA Overdier D.G., Olson E.R., Erickson B.D., Ederer M.M., Csonka L.N.;
RT "Nucleotide sequence of the transcriptional control region of the
RT osmotically regulated proU operon of Salmonella typhimurium and
RT identification of the 5' endpoint of the proU mRNA."
RL J. Bacteriol. 171:4694-4706(1989).
RN [4]
RP SEQUENCE OF 1-196 FROM N.A.
RX MEDLINE=94253018; PubMed=8195103;
RA Jordan A., Gibert I., Barbe J.;
RT "Cloning and sequencing of the genes from Salmonella typhimurium
RT encoding a new bacterial ribonucleotide reductase."
RL J. Bacteriol. 176:3420-3427(1994).
CC -1- FUNCTION: INVOLVED IN A MULTICOMPONENT BINDING-PROTEIN-DEPENDENT
CC TRANSPORT SYSTEM FOR GLYCINE BETAINES/L-PROLINE.
CC -1- SUBCELLULAR LOCATION: Inner membrane-associated.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
CC -1- SIMILARITY: Contains 2 CBS domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X52693; CAA36921.1; -.
CC EMBL; AF008828; AAL21694.1; -.
CC EMBL; M26063; AAA88621.1; -.
CC EMBL; X73226; CAA51696.1; -.
CC PIR; S05374; QREBVT.
CC HSSP; Q58663; 1G6H.
CC StyGene; SG10312; prov.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR003439; ABC_transporter.
CC InterPro; IPR000644; CBS_domain.
CC InterPro; IPR005892; Prov.
CC Pfam; PF00005; ABC_tran; 1.
CC Pfam; PF00571; CBS; 2.
CC ProDom; PD000006; ABC_transporter; 1.
CC SMART; SM00382; AAA; 1.

```

```
DR PFam: PF03799; FtsQ; 1.
KW Cell division; Septation; Transmembrane; Inner membrane;
KW Complete proteome.
FT DOMAIN 1 52 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 53 74 POTENTIAL.
FT DOMAIN 75 309 PERIPLASMIC (POTENTIAL).
FT SEQUENCE 309 AA; 34205 MW; 7DCF48DC3AF923D5 CRC64;

Query Match 78.0%; Score 32; DB 1; Length 309;
Best Local Similarity 66.7%; Pred. No. 10;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAAQRRV 9
Db 301 MLAAQEKRI 309
|||||:

RESULT 4
DHAS_ARCFU STANDARD; PRT; 343 AA.
ID DHAS_ARCFU
AC O28766;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Aspartate-semialdehyde dehydrogenase (EC 1.2.1.11) (ASA
DE dehydrogenase) (ASADH).
DE ASD OR AF1506.
GN Archaeoglobus fulgidus.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Richardson D.L., Kerlavage A.R., Gwinn M., Hickey E.K., Peterson J.D.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA Sadov P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser G.C., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic,
RT sulphate-reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
CC -!- FUNCTION: THIS ENZYME CATALYZES THE SECOND STEP IN THE COMMON
CC METABOLIC PATHWAY TO SYNTHESIZE THR AND MET FROM ASPARTIC ACID.
CC -!- CATALYTIC ACTIVITY: L-aspartate-4-semialdehyde + phosphate +
CC NADP(+) = L-4-aspartyl phosphate + NADPH.
CC -!- PATHWAY: METHIONINE BIOSYNTHESIS, THREONINE BIOSYNTHESIS.
CC -!- SIMILARITY: Belongs to the aspartate-semialdehyde dehydrogenase
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000998; AAB89738.1;
CC PIR; A69438; A69438.
CC TIGR; AF1506.
CC InterPro; IPR005676; Asp_ADH.f.
CC InterPro; IPR000319; Asp_semiadh.
CC InterPro; IPR000534; Semiadh_dh.
CC Pfam; PF01118; Semiadh_dh; 1.
CC Pfam; PF02774; Semiadh_dhc; 1.

FT DISULFID 586 597 BY SIMILARITY.
FT DISULFID 591 607 BY SIMILARITY.
FT DISULFID 609 618 BY SIMILARITY.
FT CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 194 194 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 378 378 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 515 515 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 623 623 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 751 751 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 754 754 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 900 900 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 1246 AA; 157115 MW; 255EDD7A62C025DB CRC64;

Query Match 80.5%; Score 33; DB 1; Length 1429;
Best Local Similarity 77.8%; Pred. No. 29;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAAQRRV 9
Db 1246 MLAAQEGRI 1254
|||||:

RESULT 3
FTSQ_RHIME STANDARD; PRT; 309 AA.
ID FTSQ_RHIME
AC O30993;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cell division protein ftsQ homolog.
GN FTSQ OR R02170 OR SMC01872.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=98012980; PubMed=9352931;
RA Ma X., Sun Q., Wang R., Singh G., Jonietz E.L., Margolin W.;
RA "Interactions between heterologous FtsA and FtsZ proteins at the FtsZ
RA ring.";
RL J. Bacteriol. 179:6788-6797(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Bolstad P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kias E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RA "Analysis of the chromosome sequence of the legume symbiont
RA Sinorhizobium meliloti strain 1021.";
RT Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
CC -!- FUNCTION: THIS PROTEIN MAY BE INVOLVED IN SEPTUM FORMATION (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. INNER MEMBRANE
CC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE FTSQ FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF024660; AAC45822.1; ALT_INIT.
CC EMBL; AL501789; CAC46749.1;
CC InterPro; IPR005548; FtsQ.
```

```

RESULT 2
L112_CAEEL STANDARD; PRT; 1429 AA.
AC P14585;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Lin-12 protein precursor.
GN Lin-12 OR R107.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN-Bristol N2;
RC MEDLINE=88334747; PubMed=3419531;
RX Yochem J., Weston K., Greenwald I.;
RA "The Caenorhabditis elegans lin-12 gene encodes a transmembrane
RT protein with overall similarity to Drosophila Notch.";
RL Nature 335:547-550(1988).
RN [2]
RN SEQUENCE FROM N.A.
RP STRAIN-Bristol N2;
RC MEDLINE=94150718; PubMed=7906398;
RX Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkneen R.,
RA Sims M., Smaldon N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RL elegans.";
CC -1- FUNCTION: LIN-12 IS INVOLVED IN SEVERAL CELL FATES DECISIONS THAT
CC REQUIRES CELL-CELL INTERACTIONS. IT IS POSSIBLE THAT LIN-12
CC ENCODES A MEMBRANE-BOUND RECEPTOR FOR A SIGNAL THAT ENABLES
CC EXPRESSION OF THE VENTRAL UTERINE PRECURSOR CELL FATE.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: HIGH, TO C.ELEGANS GLP-1.
CC -1- SIMILARITY: Contains 13 EGF-like domains.
CC -1- SIMILARITY: Contains 3 lin/Notch repeats.
CC -1- SIMILARITY: Contains 5 ANK repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR DR EMBL; M12069; AAA70191.1; -;
DR DR EMBL; Z14092; CAA78474.1; -;
DR DR PIR; S06434; S06434.
DR DR HSSP; P00740; IEDM.
DR DR WormFep; R107.8; CE00274.
DR DR InterPro; IPR002110; ANK.
DR DR InterPro; IPR000152; Asx_hydroxyl.
DR DR InterPro; IPR000742; EGF_2.
DR DR InterPro; IPR001881; EGF_Ca.
DR DR InterPro; IPR006209; EGF_like.
DR DR InterPro; IPR002049; Laminin_EGF.
DR DR InterPro; IPR000800; Notch.
DR DR Pfam; PF00023; ank; 6.
DR DR Pfam; PF00008; EGF; 13.
DR DR Pfam; PF00066; notch; 3.
DR DR PRINTS; PR00011; EGF_LAMININ.
DR DR PRINTS; PR01452; NOTCH.
DR DR SMART; SM00248; ANK; 6.
DR DR SMART; SM00179; EGF_CA; 2.
DR DR SMART; SM00004; NL; 2.
DR DR PROSITE; PS00088; ANK_REPEAT; 3.
DR DR PROSITE; PS0297; ANK_REPEAT_REGION; 1.
DR DR PROSITE; PS00010; ASX_HYDROXYL; 3.
DR DR PROSITE; PS00022; EGF_1; 12.
DR DR PROSITE; PS01186; EGF_2; 11.
DR DR PROSITE; PS01187; EGF_CA; 2.
KW Differentiation; Repeat; ANK repeat; EGF-like domain; Transmembrane;
KW Glycoprotein; Signal.
FT SIGNAL 1 15 POTENTIAL.
FT CHAIN 16 1429 LIN-12 PROTEIN.
FT DOMAIN 16 908 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 909 931 POTENTIAL.
FT DOMAIN 932 1429 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 20 61 EGF-LIKE 1.
FT DOMAIN 114 150 EGF-LIKE 2.
FT DOMAIN 152 190 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 201 246 EGF-LIKE 4.
FT DOMAIN 250 285 EGF-LIKE 5.
FT DOMAIN 287 323 EGF-LIKE 6.
FT DOMAIN 323 363 EGF-LIKE 7.
FT DOMAIN 365 402 EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 404 441 EGF-LIKE 9.
FT DOMAIN 449 492 EGF-LIKE 10.
FT DOMAIN 503 541 EGF-LIKE 11.
FT DOMAIN 543 579 EGF-LIKE 12.
FT DOMAIN 582 619 EGF-LIKE 13.
FT REPEAT 635 669 LIN/NOTCH 1.
FT REPEAT 670 710 LIN/NOTCH 2.
FT REPEAT 711 750 LIN/NOTCH 3.
FT REPEAT 1093 1122 ANK 1.
FT REPEAT 1126 1158 ANK 2.
FT REPEAT 1162 1194 ANK 3.
FT REPEAT 1206 1236 ANK 4.
FT REPEAT 1240 1269 ANK 5.
FT DISULFID 24 35 BY SIMILARITY.
FT DISULFID 29 49 BY SIMILARITY.
FT DISULFID 51 60 BY SIMILARITY.
FT DISULFID 118 129 BY SIMILARITY.
FT DISULFID 123 138 BY SIMILARITY.
FT DISULFID 140 149 BY SIMILARITY.
FT DISULFID 156 169 BY SIMILARITY.
FT DISULFID 163 178 BY SIMILARITY.
FT DISULFID 180 189 BY SIMILARITY.
FT DISULFID 205 227 BY SIMILARITY.
FT DISULFID 221 234 BY SIMILARITY.
FT DISULFID 236 245 BY SIMILARITY.
FT DISULFID 254 264 BY SIMILARITY.
FT DISULFID 259 273 BY SIMILARITY.
FT DISULFID 275 284 BY SIMILARITY.
FT DISULFID 291 302 BY SIMILARITY.
FT DISULFID 296 311 BY SIMILARITY.
FT DISULFID 313 322 BY SIMILARITY.
FT DISULFID 327 339 BY SIMILARITY.
FT DISULFID 334 351 BY SIMILARITY.
FT DISULFID 353 362 BY SIMILARITY.
FT DISULFID 369 381 BY SIMILARITY.
FT DISULFID 375 390 BY SIMILARITY.
FT DISULFID 392 401 BY SIMILARITY.
FT DISULFID 408 419 BY SIMILARITY.
FT DISULFID 413 429 BY SIMILARITY.
FT DISULFID 431 440 BY SIMILARITY.
FT DISULFID 507 518 BY SIMILARITY.
FT DISULFID 512 529 BY SIMILARITY.
FT DISULFID 531 540 BY SIMILARITY.
FT DISULFID 547 558 BY SIMILARITY.
FT DISULFID 552 567 BY SIMILARITY.
FT DISULFID 569 578 BY SIMILARITY.

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:05:57 ; Search time 2.62025 Seconds  
(without alignments)  
161.527 Million cell updates/sec

Title: US-09-807-512-26  
Perfect score: 41  
Sequence: 1 MLAAQERRV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues  
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	33	80.5	482	1 MURC_PASMU	P57818 pasteurella
2	33	80.5	1429	1 L112_CAEEL	P14585 caenorhabdi
3	32	78.0	309	1 FTSQ_RHIME	O30993 rhizobium m
4	31	75.6	343	1 DHAS_ARCFU	O28766 archaeglob
5	30	73.2	376	1 Y592_METWA	O892b2 methanosarc
6	30	73.2	400	1 PROV_SALTY	P17328 salmonella
7	29	70.7	140	1 YF98_METTH	O27635 mechanobact
8	29	70.7	258	1 CYSH_VIBVU	O8CwK6 vibrio vuln
9	29	70.7	267	1 MSX2_HUMAN	P35548 homo sapien
10	29	70.7	292	1 HCHA_STAAM	O99w58 staphylococ
11	29	70.7	292	1 HCHA_STAAM	O8bxy2 staphylococ
12	29	70.7	400	1 CYH2_HUMAN	O99418 homo sapien
13	29	70.7	400	1 CYH2_MOUSE	P97695 mus musculu
14	29	70.7	494	1 CORQ_MYCTU	O53677 mycobacteri
15	29	70.7	530	1 YK03_YEAST	P36119 saccharomyc
16	29	70.7	582	1 HS60_SCHPO	O09864 schizosacch
17	29	70.7	860	1 KDPD_MYCTU	P96372 mycobacteri
18	29	70.7	3164	1 TEGU_HSV11	P10220 herpes simp
19	29	70.7	4377	1 ANK3_HUMAN	Q12955 homo sapien
20	28	68.3	183	1 Y584_CHLTR	O84588 chlamydia t
21	28	68.3	183	1 Y873_CHLMU	O9pjf6 chlamydia m
22	28	68.3	189	1 INA6_HUMAN	P05013 homo sapien
23	28	68.3	189	1 INAD_HUMAN	P01570 homo sapien
24	28	68.3	190	1 MOBA_RHOCA	O9x7k0 rhodobacter
25	28	68.3	193	1 MOBA_DEIRA	Q9rwm1 deinococcus
26	28	68.3	205	1 CEAB_ECOLI	P09882 escherichia
27	28	68.3	265	1 FABI_PSEAE	O9zfe4 pseudomonas
28	28	68.3	399	1 CYH3_MOUSE	O08967 mus musculu
29	28	68.3	400	1 CYH3_HUMAN	O43739 homo sapien
30	28	68.3	400	1 CYH3_RAT	P97696 rattus norv
31	28	68.3	403	1 PRS8_SCHPO	P41836 schizosacch
32	28	68.3	414	1 PRS8_NAEFO	Q25544 naegleria f
33	28	68.3	427	1 DHT2_SHEEP	P50168 ovnis aries

34	28	68.3	427	1 SYS_BRUME	Q8ygs6 brucella me
35	28	68.3	427	1 SYS_BRUSU	Q8g139 brucella su
36	28	68.3	534	1 TYR1_AMAME	P55027 ambystoma m
37	28	68.3	573	1 AMH2_HUMAN	Q16671 homo sapien
38	28	68.3	581	1 CEA2_ECOLI	P04419 escherichia
39	28	68.3	582	1 CEA9_ECOLI	P09883 escherichia
40	28	68.3	637	1 PBPA_STROK	Q00573 streptococc
41	28	68.3	1053	1 SAL4_HUMAN	Q9ujq4 homo sapien
42	28	68.3	1158	1 KCH2_CANFA	Q9Lsz3 canis famil
43	28	68.3	1159	1 KCH2_HUMAN	Q12809 homo sapien
44	28	68.3	1161	1 KCH2_RABIT	Q8wny2 oryctolagus
45	28	68.3	1162	1 KCH2_MOUSE	O35219 mus musculu

ALIGNMENTS

RESULT 1  
MURC\_PASMU STANDARD; PRT; 482 AA.  
AC P57818;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE UDP-N-acetylmuramate-L-alanine ligase (EC 6.3.2.8) (UDP-N-  
DE acetylmuramoyl-L-alanine synthetase).  
GN MURC OR PM0143.  
OS Pasteurella multocida.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
OC Pasteurellaceae; Pasteurella.  
OX NCBI\_TaxID=747;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21145866; PubMed=11248100;  
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;  
RT "Complete genomic sequence of Pasteurella multocida Pn70.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).  
CC -!- FUNCTION: CELL WALL FORMATION  
CC -!- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl + L-alanine = ADP +  
CC phosphate + UDP-N-acetylmuramoyl-L-alanine.  
CC -!- PATHWAY: Peptidoglycan biosynthesis.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
CC -!- SIMILARITY: Belongs to the murCDEF family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AE006049; AAK02227.1; -;  
DR HAMAP; MF\_00046; -; 1.  
DR InterPro; IPR000713; Mur\_ligase.  
DR InterPro; IPR004101; Mur\_ligase\_C.  
DR InterPro; IPR005758; MurC.  
DR Pfam; PF01225; Mur\_ligase; 1.  
DR Pfam; PF02875; Mur\_ligase; 1.  
DR TIGRFAMs; TIGR01082; murC; 1.  
KW Ligase; ATP-binding; Cell division; Cell wall;  
KW Peptidoglycan synthesis; Complete proteome.  
FT NP\_BIND 131 137 ATP (POTENTIAL).  
SQ SEQUENCE 482 AA; 52731 MW; 98667510E7A2D503 CRC64;  
Query Match 80.5%; Score 33; DB 1; Length 482;  
Best Local Similarity 66.7%; Pred. No. 9.7;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MLAAQERRV 9  
DB 100 VVAAQERRI 108

C;Genetics:  
A;Gene: SCQEDB:SC4A2.16c

Query Match 70.7%; Score 29; DB 2; Length 118;  
Best Local Similarity 85.7%; Pred. No. 46;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAAQERR 8  
Db 46 LSAQERR 52

## RESULT 12

B95873  
hypothetical protein [imported] - Sinorhizobium meliloti (strain 1021) megaplasmid pSymB  
C;Species: Sinorhizobium meliloti  
C;Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001  
C;Accession: B95873  
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan  
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
A;Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing endo  
A;Reference number: A95842; MUID:21396508; PMID:11481431  
A;Accession: B95873  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-130 <KUR>  
A;Cross-references: GB:AL501985; PIDN:CAC48650.1; PID:g15140122; GSPDB:GN00167  
A;Experimental source: strain 1021, megaplasmid pSymB  
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,  
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;  
L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.  
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
A;Reference number: A96039; MUID:21368234; PMID:11474104  
C;Contents: annotation  
C;Genetics:  
A;Gene: Smb20260  
A;Genome: plasmid

Query Match 70.7%; Score 29; DB 2; Length 130;  
Best Local Similarity 75.0%; Pred. No. 51;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAAQERR 9  
Db 23 LAARDRRV 30

## RESULT 13

C69080  
conserved hypothetical protein MTH1598 - Methanobacterium thermoautotrophicum (strain De  
C;Species: Methanobacterium thermoautotrophicum  
C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 08-Oct-1999  
C;Accession: C69080  
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;  
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.  
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997  
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct  
A;Reference number: A69000; MUID:98037514; PMID:9371463  
A;Accession: C69080  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-140 <MTH>  
A;Cross-references: GB:AE000919; GB:AE000666; NID:g2622717; PIDN:AAB86071.1; PID:g262272  
A;Experimental source: strain Delta H  
C;Genetics:  
A;Gene: MTH1598  
C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0024

Query Match 70.7%; Score 29; DB 2; Length 140;

Best Local Similarity 85.7%; Pred. No. 55;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAOERRV 9  
Db 45 AAERRV 51

## RESULT 14

B83975  
hypothetical protein BH2602 [imported] - Bacillus halodurans (strain C-125)  
C;Species: Bacillus halodurans  
C;Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
C;Accession: B83975  
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, N.; Masui, R.; Fujii, F.; H  
Nucleic Acids Res. 28, 4317-4331, 2000  
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a  
A;Reference number: A83650; MUID:20512582; PMID:11058132  
A;Accession: B83975  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-168 <STO>  
A;Cross-references: GB:AP001516; GB:BA000004; NID:g10175192; PIDN:BAB06321.1; GSPDB:G  
A;Experimental source: strain C-125  
C;Genetics:  
A;Gene: BH2602

Query Match 70.7%; Score 29; DB 2; Length 168;  
Best Local Similarity 75.0%; Pred. No. 66;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAAQERR 8  
Db 36 LLAQERR 43

## RESULT 15

T34834  
probable transferase - Streptomyces coelicolor (fragment)  
C;Species: Streptomyces coelicolor  
C;Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 03-Nov-2000  
C;Accession: T34834  
R;Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A  
submitted to the EMBL Data Library, February 1999  
A;Reference number: Z21559  
A;Accession: T34834  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-187 <OLI>  
A;Cross-references: EMBL:AL035478; PIDN:CAB36588.2; GSPDB:GN000070; SCQEDB:SC265.01  
A;Experimental source: strain A3(2)  
C;Genetics:  
C;Superfamily: Rhizobium nodulation protein nodU

Query Match 70.7%; Score 29; DB 2; Length 187;  
Best Local Similarity 87.5%; Pred. No. 73;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAAQERR 8  
Db 129 MAAQERR 136

Search completed: October 7, 2003, 13:32:41  
Job time : 6.06962 secs

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MUID:98295987; PMID:9634230  
A:Accession: A70842  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-394 <COL>  
A:Cross-references: GB:AL021841; GB:AL123456; NID:g3261517; PIDN:CAA17078.1; PID:g289421  
A:Experimental source: strain H37Rv  
C:Genetics:  
A:Gene: amiB  
C:Superfamily: hippurate hydrolase

Query Match 73.2%; Score 30; DB 2; Length 394;  
Best Local Similarity 75.0%; Pred. No. 94;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAQERR 8  
:|||||  
Db 384 VLAQQR 391

RESULT 8  
QREBVT  
N:Alternate names: nucleotide-binding protein prov - Salmonella typhimurium  
C:Species: Salmonella typhimurium  
C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 16-Jul-1999  
C:Accession: S05374; B45917; S34273  
R:Stirling, D.A.; Hulton, C.S.J.; Waddell, L.; Park, S.F.; Stewart, G.S.A.B.; Booth, I.F.  
Mol. Microbiol. 3, 1025-1038, 1989  
A:Title: Molecular characterization of the proU loci of Salmonella typhimurium and Esche  
A:Reference number: S05374; MUID:90113884; PMID:2691838  
A:Accession: S05374  
A:Molecule type: DNA  
A:Residues: 1-400 <STI>  
A:Cross-references: GB:X52693; NID:g47829; PIDN:CAA36921.1; PID:g47831  
R:Overdier, D.G.; Olson, E.R.; Erickson, B.D.; Ederer, M.M.; Csonka, L.N.  
J. Bacteriol. 171, 4694-4706, 1989  
A:Title: Nucleotide sequence of the transcriptional control region of the osmotically re  
A:Reference number: A45917; MUID:89359099; PMID:2548994  
A:Accession: B45917  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-157, 'A', 159-290 <OVE>  
A:Cross-references: GB:M26063; NID:g154278; PIDN:AAA88621.1; PID:g552022  
R:Jordan Valles, A.  
Submitted to the EMBL Data Library, June 1993  
A:Reference number: S34271  
A:Accession: S34273  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-157, 'A', 159-196 <JOR>  
A:Cross-references: EMBL:X73226  
C:Comment: This protein is the hydrophilic nucleotide-binding component of the binding p  
C:Genetics:  
A:Gene: prov  
A:Map position: 57 min  
C:Superfamily: glycine betaine/proline transport protein prov; ATP-binding cassette hom  
C:Keywords: ATP; binding protein-dependent transport system; glycine betaine transport;  
F:44-241/Domain: ATP-binding cassette homology <ABC>  
F:161-68/Region: nucleotide-binding motif A (P-loop)  
F:185-189/Region: nucleotide-binding motif B  
F:286-334/Domain: CBS homology <CBS1>  
F:346-393/Domain: CBS homology <CBS2>

Query Match 73.2%; Score 30; DB 1; Length 400;  
Best Local Similarity 85.7%; Pred. No. 95;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAAQERR 8

Db 136 IAAQERR 142  
:|||||  
RESULT 9  
T25782  
hypothetical protein F47D2.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T25782  
R:Bradshaw, H.  
Submitted to the EMBL Data Library, July 1996  
A:Description: The sequence of C. elegans cosmid F47D2.  
A:Reference number: Z20085  
A:Accession: T25782  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-1047 <BRA>  
A:Cross-references: EMBL:U64846; PIDN:AAB04866.1; GSPDB:GN00023; CESP:F47D2.2  
A:Experimental source: strain Bristol N2; clone F47D2  
C:Genetics:  
A:Gene: CESP:F47D2.2  
A:Map position: 5  
A:Introns: 159/2; 230/3; 446/3; 779/1; 840/3

Query Match 73.2%; Score 30; DB 2; Length 1047;  
Best Local Similarity 77.8%; Pred. No. 2.5e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAAQERR 9  
:|||||  
Db 363 MAAQERR 371

RESULT 10  
JS0654  
hypothetical 8.2K protein (kan 5' region) - Streptomyces griseus (strain SS-1198Pr)  
C:Species: Streptomyces griseus  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 18-Nov-1994  
C:Accession: JS0654  
R:Shikawa, J.; Hotta, K.  
Gene 108, 127-132, 1991  
A:Title: Nucleotide sequence and transcriptional start point of the kan gene encoding  
A:Reference number: JS0652; MUID:92104494; PMID:1761222  
A:Accession: JS0654  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-72 <ISH>

Query Match 70.7%; Score 29; DB 2; Length 72;  
Best Local Similarity 85.7%; Pred. No. 28;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAAQERR 8  
:|||||  
Db 46 LAAQRR 52

RESULT 11  
T34991  
hypothetical protein SC4A2.16c SC4A2.16c - Streptomyces coelicolor  
C:Species: Streptomyces coelicolor  
C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999  
C:Accession: T34991  
R:Oliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
Submitted to the EMBL Data Library, August 1998  
A:Reference number: Z21550  
A:Accession: T34991  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-118 <OLI>  
A:Cross-references: EMBL:AL031182; PIDN:CAA20168.1; GSPDB:GN00070; SCOEDB:SC4A2.16c  
A:Experimental source: strain A3(2)

Smith, H.O.; Woese, C.R.; Venter, J.C.  
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae  
A:Reference number: A69250; MUID:98049343; PMID:9389475  
A:Accession: A69438  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-343 <KLE>  
A:Cross-references: GB:AE000098; GB:AE000782; NID:g2689321; PIDN:AAB89738.1; PID:g264905  
C:Superfamily: aspartate-semialdehyde dehydrogenase

Query Match 75.6%; Score 31; DB 2; Length 343;  
Best Local Similarity 87.5%; Pred. No. 49;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAAQERRV 9  
Db 33 LAASERV 40  
||| |||||

RESULT 3  
S62692  
L-amino acid oxidase - *Synechococcus* sp. (PCC 6301)  
C:Species: *Synechococcus* sp.  
A:Variety: PCC 6301  
C:Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 29-Aug-1997  
C:Accession: S62692; S62687  
R:Bockholt, R.; Masepohl, B.; Kruft, V.; Wittmann-Liebold, B.; Pistorius, E.K.  
Biochim. Biophys. Acta 1264, 289-293, 1995  
A:Title: Partial amino acid sequence of an L-amino acid oxidase from the cyanobacterium  
A:Reference number: S62687; MUID:96138550; PMID:8547316  
A:Accession: S62692  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-355 <BOC>  
A:Cross-references: EMBL:Z48565  
A:Accession: S62687  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 52-68 <BO2>

Query Match 75.6%; Score 31; DB 2; Length 355;  
Best Local Similarity 87.5%; Pred. No. 51;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAAQERRV 9  
Db 95 LAPQERRV 102  
||| |||||

RESULT 4  
A87645  
efflux system protein [imported] - *Caulobacter crescentus*  
C:Species: *Caulobacter crescentus*  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
C:Accession: A87645  
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of *Caulobacter crescentus*.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: A87645  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-500 <STO>  
A:Cross-references: GB:AE005673; NID:g13424871; PIDN:AAK25157.1; GSPDB:GN00148  
C:Genetics:  
A:Gene: CC3195

Query Match 75.6%; Score 31; DB 2; Length 500;  
Best Local Similarity 87.5%; Pred. No. 72;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAAQERRV 9  
Db 414 LAADERRV 421  
||| |||||

## RESULT 5

A71105  
hypothetical protein PH0610 - *Pyrococcus horikoshii*  
C:Species: *Pyrococcus horikoshii*

C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 28-Jul-2000

C:Accession: A71105

R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Hailkawa, Y.; Hino, Y.; Yamamoto, S.; Se  
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Ogu  
DNA Res. 5, 55-76, 1998

A:Title: Complete sequence and gene organization of the genome of a hyper-thermophil

A:Reference number: A71000; MUID:98344137; PMID:9679194

A:Accession: A71105

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-165 <RAW>

A:Cross-references: GB:AP000003; NID:g3236130; PIDN:BAA29699.1; PID:g3257016

A:Experimental source: strain OT3

A:Note: this accession replaces an interim accession for a sequence replaced by GenB

C:Genetics:

A:Gene: PH0610

C:Superfamily: *Pyrococcus horikoshii* hypothetical protein PH1177

Query Match 73.2%; Score 30; DB 2; Length 165;

Best Local Similarity 75.0%; Pred. No. 39;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAAQERRV 9

Db 144 LAPQERRI 151

||| |||||

## RESULT 6

C75056

hypothetical protein PAB0956 - *Pyrococcus abyssi* (strain Orsay)

C:Species: *Pyrococcus abyssi*

C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 28-Jul-2000

C:Accession: C75056

R:anonymous, Genoscope

submitted to the EMBL Data Library, July 1999

A:Description: *Pyrococcus abyssi* genome sequence: insights into archaeal chromosome s

A:Reference number: A75001

A:Accession: C75056

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-168 <RAW>

A:Cross-references: GB:AJ248287; GB:AL096836; NID:g5458657; PIDN:CAB50344.1; PID:g545

A:Experimental source: strain Orsay

C:Genetics:

A:Gene: PAB0956

C:Superfamily: *Pyrococcus horikoshii* hypothetical protein PH1177

Query Match 73.2%; Score 30; DB 2; Length 168;

Best Local Similarity 75.0%; Pred. No. 40;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAAQERRV 9

Db 144 LAPQERRI 151

||| |||||

## RESULT 7

A70842

probable *amiB* protein - *Mycobacterium tuberculosis* (strain H37Rv)

C:Species: *Mycobacterium tuberculosis*

C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000

C:Accession: A70842

R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon

; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:19:12 ; Search time 5.06962 Seconds  
(without alignments)  
170.726 Million cell updates/sec

Title: US-09-807-512-26

Perfect score: 41

Sequence: 1 MAAQERRV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_76:\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	80.5	1429	2 S06434	homeotic protein 1
2	31	75.6	343	2 A69438	aspartate-semialde
3	31	75.6	355	2 S62692	L-amino acid oxida
4	31	75.6	500	2 A87645	efflux system prot
5	30	73.2	165	2 A71105	hypothetical prote
6	30	73.2	168	2 C75056	hypothetical prote
7	30	73.2	394	2 A70842	probable amib prot
8	30	73.2	400	1 QREBVT	glycine betaine/pr
9	30	73.2	1047	2 T25782	hypothetical prote
10	29	70.7	72	2 JS0654	hypothetical 8.2K
11	29	70.7	118	2 T34991	hypothetical prote
12	29	70.7	130	2 B95873	hypothetical prote
13	29	70.7	140	2 C69080	conserved hypothet
14	29	70.7	168	2 B83975	hypothetical prote
15	29	70.7	187	2 T34834	probable transfera
16	29	70.7	197	2 H75563	transcription regu
17	29	70.7	204	2 C70506	hypothetical prote
18	29	70.7	235	2 E75375	conserved hypothet
19	29	70.7	267	2 A49068	craniosynotosis-as
20	29	70.7	292	2 A89823	conserved hypothet
21	29	70.7	306	2 D87531	glycosyl transfera
22	29	70.7	328	2 S34388	NADH2 dehydrogenas
23	29	70.7	332	2 S34389	NADH2 dehydrogenas
24	29	70.7	406	2 T30141	hypothetical prote
25	29	70.7	494	2 C70940	probable cobQ prot
26	29	70.7	530	2 S38092	hypothetical prote
27	29	70.7	540	2 T49184	hypothetical prote
28	29	70.7	557	2 G81941	probable DNA repair
29	29	70.7	557	2 D81165	DNA repair protein

30	29	70.7	569	2 T47358	hypothetical prote
31	29	70.7	582	2 S62535	heat shock protein
32	29	70.7	582	2 T43369	heat-shock protein
33	29	70.7	656	2 AB1843	hypothetical prote
34	29	70.7	675	2 T32299	hypothetical prote
35	29	70.7	860	2 G70623	probable sensor pr
36	29	70.7	1034	2 T32297	hypothetical prote
37	29	70.7	1154	2 AI3431	chromosome segrega
38	29	70.7	1461	2 B70588	probable polyketid
39	29	70.7	1765	2 T42714	ankyrin 3, splice
40	29	70.7	1940	2 T42715	ankyrin 3, splice
41	29	70.7	1943	2 T42713	ankyrin 3, splice
42	29	70.7	1961	2 T42716	ankyrin 3, splice
43	29	70.7	3164	1 WMBEH6	UL36 protein - hum
44	29	70.7	3429	2 T13853	hypothetical prote
45	29	70.7	4377	2 A55575	ankyrin 3, long sp

#### ALIGNMENTS

##### RESULT 1

S06434

homeotic protein lin-12 precursor - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Jan-1993 #sequence\_revision 29-Jan-1993 #text\_change 20-Sep-1999

C:Accession: S06434; A24769

R:Yochem, J.; Weston, K.; Greenwald, I.

Nature 335, 547-550, 1988

A:Title: The Caenorhabditis elegans lin-12 gene encodes a transmembrane protein with  
A:Reference number: S06434; MUID:88334747; PMID:3419531

A:Accession: S06434

A:Molecule type: DNA

A:Residues: 1-1429 <YOC>

A:Cross-references: EMBL:M12069; NID:gl56357; PIDN:AAA70191.1; PID:gl56358

R:Greenwald, I.

Cell 43, 583-590, 1985

A:Reference number: A24769; MUID:86079540; PMID:3000611

A:Accession: A24769

A:Molecule type: DNA

A:Residues: 173-712 <GRE>

C:Genetics:

A:Introns: 50/2; 90/1; 109/1; 172/3; 545/1; 589/2; 632/2; 1273/3; 1389/3

C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homol

C:Keywords: glycoprotein; transmembrane protein

F:254-284/Domain: EGF homology <EGF1>

F:507-540/Domain: EGF homology <EGF>

F:547-578/Domain: EGF homology <EGF2>

F:909-931/Domain: transmembrane #status predicted <TMM>

F:1093-1125/Domain: ankyrin repeat homology <AN1>

F:1206-1238/Domain: ankyrin repeat homology <AN2>

F:1240-1272/Domain: ankyrin repeat homology <AN3>

Query Match 80.5%; Score 33; DB 2; Length 1429;

Best Local Similarity 77.8%; Pred. No. 75;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAAQERRV 9

|||||

DB 1246 MAAQEGRI 1254

##### RESULT 2

A69438

aspartate-semialdehyde dehydrogenase (asd) homolog - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus

C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 04-Nov-2002

C:Accession: A69438

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod

.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes,



XX SQ Sequence 3541 AA;  
 Query Match 100.0%; Score 41; DB 23; Length 3541;  
 Best Local Similarity 100.0%; Pred. No. 25;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 M LAAQERRV 9  
 Db 184 M LAAQERRV 192

RESULT 14  
 AAY05982  
 ID AAY05982 standard; Peptide; 10 AA.  
 XX AC AAY05982;  
 XX DT 16-AUG-1999 (first entry)  
 XX DE Human cancer antigen NY ESO-1/CAG-3 ORF2 cancer peptide.  
 XX NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;  
 KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
 KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
 KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
 KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
 KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
 KW vaccine.  
 XX OS Homo sapiens.  
 XX PN W09918206-A2.  
 XX PD 15-APR-1999.  
 XX PF 21-SEP-1998; 98WO-US19609.  
 XX PR 08-OCT-1997; 97US-0061428.  
 XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX PI Rosenberg SA, Wang RF;  
 XX WPI; 1999-277270/23.  
 XX DR N-PSDB; AAX58601.  
 XX Cancer antigen NY ESO1/CAG-3  
 XX Claim 26; Page 65; 88pp; English.  
 XX The present sequence represents a cancer peptide that is based on  
 CC amino acid residues 19-27 of human ESO-1/CAG-3 (or CAG-3) ORF2  
 CC (see AAY05966), a new and potent tumour antigen capable of eliciting  
 CC an antigen specific immune response by T cells. Cancer peptides  
 CC derived from CAG-3 ORF2, CAG-3 ORF1 (see AAY05965), portions of them  
 CC and their variants (see AAY05967-87), are useful as cancer vaccines  
 CC that protect against cancer. The invention provides: vectors and  
 CC host cells (also useful as vaccines); a method of diagnosis of  
 CC cancer or precancer; a transgenic animal; antisense oligonucleotides  
 CC that inhibit expression of the cancer peptide or tumour antigen;  
 CC antibodies reacting with a CAG-3 cancer peptide, useful in  
 CC diagnostic and detection assays; and methods for preventing or  
 CC inhibiting cancer by administering a cancer peptide, with or without  
 CC an HLA molecule. The cancer peptides form part of, or are derived  
 CC from, cancers such as primary or metastatic melanoma, thymoma,  
 CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine  
 CC cancer, cervical cancer, bladder cancer, kidney cancer and  
 CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and  
 CC thyroid cancers. Melanoma is treated by inducing cancer-specific T  
 CC cells in vitro for subsequent return to a patient.  
 XX SQ Sequence 10 AA;

Query Match 87.8%; Score 36; DB 20; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.74;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAAQERRV 9  
 Db 1 LAAQERRV 8

RESULT 15  
 AAB31332  
 ID AAB31332 standard; peptide; 10 AA.  
 XX AC AAB31332;  
 XX DT 20-APR-2001 (first entry)  
 XX DE Exemplary antigen characteristic of tumours and derived from NY-ESO-1.  
 XX MAGE-A1; HLA; human leukocyte antigen; CD4+ T lymphocyte; cancer;  
 KW MAGE-A1 HLA class II-binding protein; vaccine.  
 XX OS Homo sapiens.  
 XX PN W0200078806-A1.  
 XX PD 28-DEC-2000.  
 XX PF 14-JUN-2000; 2000WO-US16287.  
 XX PR 18-JUN-1999; 9905-0336091.  
 XX PA (LUDW-) LUDWIG INST CANCER RES.  
 XX PI Van Snick J, Lethe B, Chau P, Boon-Falleur T, Van Der Bruggen P;  
 XX WPI; 2001-102698/11.  
 XX Novel MAGE-A1 human leukocyte antigen class II peptides which bind to  
 PT and are presented to the class II molecules, useful for inducing immune  
 PT response and treating cancers characterized by expression of MAGE-A1 -  
 XX Disclosure; Page 32; 78pp; English.  
 XX AAB31302-59 represent exemplary antigens which are characteristic  
 CC of tumours. They can be used to enhance the immune response of vaccines  
 CC comprising peptides derived from human MAGE-A1 HLA (human leukocyte  
 CC antigen) class II-binding protein. Peptides derived from the MAGE-A1 HLA  
 CC binding protein stimulate the activity and proliferation of CD4+ T  
 CC lymphocytes. The MAGE-A1 HLA binding protein is useful as a diagnostic  
 CC agent for diagnosing a disorder characterized by expression of MAGE-A1.  
 CC The protein is used for treating a disorder characterized by expression  
 CC of MAGE-A1 such as cancers e.g. melanoma, squamous cell carcinomas,  
 CC colorectal carcinomas, osteosarcomas, and lymphocytic leukemias.  
 CC Peptides derived from the MAGE-A1 HLA binding protein are useful in the  
 CC production of anti-tumour vaccines.  
 XX SQ Sequence 10 AA;

Query Match 87.8%; Score 36; DB 22; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.74;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAAQERRV 9  
 Db 1 LAAQERRV 8

Search completed: October 7, 2003, 13:25:33  
 Job time : 14.9241 secs

XX DT 31-JUL-2000 (first entry)

XX DE Human CTL-recognised Antigen on MELanoma (CAMEL) protein.

XX KW CAMEL; CTL-recognised Antigen on MELanoma; cytotoxic T lymphocyte; CTL;

XX KW tumour-associated antigen; LAGE-1; NY-ESO-1; anticancer; melanoma;

XX KW cancer; immunotherapy; immunogenic peptide; immune response.

XX OS Homo sapiens.

XX FH Key

XX FT Peptide

XX FT 1..11

XX FT /label= CTL epitope

XX FT /note= "Immunogenic peptide with ability to elicit a

XX FT CTL response"

XX FT Peptide

XX FT 2..10

XX FT /label= Immunogenic\_peptide

XX FT /note= "Specific for HLA-A3"

XX FT Peptide

XX FT 2..11

XX FT /label= CTL epitope

XX FT /note= "Immunogenic peptide with ability to elicit a

XX FT CTL response"

XX FT Peptide

XX FT 10..18

XX FT /label= CAMEL\_10

XX FT /note= "Specific for HLA-A2"

XX FT Peptide

XX FT 16..25

XX FT /label= CAMEL\_16

XX FT /note= "Specific for HLA-A2"

XX FT Peptide

XX FT 17..25

XX FT /label= CAMEL\_17

XX FT /note= "Specific for HLA-A2"

XX FT Peptide

XX FT 51..59

XX FT /label= Immunogenic\_peptide

XX FT /note= "Specific for HLA-A3101"

XX FT Peptide

XX FT 101..109

XX FT /label= Immunogenic\_peptide

XX FT WO200023584-A1.

XX PN 27-APR-2000.

XX PD 99WO-EP07832.

XX PF 15-OCT-1999; 99WO-EP07832.

XX PR 16-OCT-1998; 98EP-0119583.

XX PR (BOEH ) BOEHRINGER INGELHEIM INT GMBH.

XX PA (UYHO-) UNIV HOSPITAL LEIDEN.

XX PI Schrier PI, Aarnoudse CA, Heider K, Klade C;

XX WPI; 2000-339685/29.

XX DR N-PSDB; AAD00149.

XX TT tumor-associated antigen useful for cancer immunotherapy is encoded by

XX FT the open reading frame of LAGE-1 (a tumor-specific antigen) CDNA -

XX PS Claim 1; Page 55; 73pp; English.

XX CC The present protein sequence is the human tumour-associated antigen CAMEL

XX CC (Cytotoxic T lymphocytes (CTL)-recognised Antigen on MELanoma). CAMEL

XX CC protein is encoded by the LAGE-1 gene, a tumour-specific antigen. It is

XX CC different from the LAGE-1 protein, since it is translated from a

XX CC different open reading frame (ORF-1). It shows strong homology with

XX CC NY-ESO-1, a melanoma specific tumour antigen. The tumour-associated

XX CC antigen displayed on melanoma cells is recognised by cytotoxic T

XX CC lymphocytes. CAMEL is expressed in tumour cell lines, tumour tissues

XX CC (e.g. breast and lung) and in restricted number of healthy tissues. This

XX CC sequence has anticancer activity. CAMEL tumour antigen and immunogenic

XX CC peptides derived from it are useful for cancer immunotherapy. They have

XX CC the potential to induce an immune response, by eliciting a CTL response.

XX CC The DNA molecule is used to construct recombinant or fusion proteins.

SQ Sequence 109 AA;

Query Match 100.0%; Score 41; DB 21; Length 109;

Best Local Similarity 100.0%; Pred. NO. 0.79;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLLAAQERRV 9

Db 17 MLLAAQERRV 25

|||||||

RESULT 13

AAU85130

ID AAU85130 standard; Protein; 3541 AA.

XX AC AAU85130;

XX DT 08-MAY-2002 (first entry)

XX DE Human melanoma specific savine.

XX KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;

XX KW viral infection; human immunodeficiency virus; melanoma;

XX KW bacterial infection; Salmonella; Legionella; parasitic infection;

XX KW Trypanosoma; Toxoplasma; Giardia.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO200190197-A1.

XX PD 29-NOV-2001.

XX PF 25-MAY-2001; 2001WO-AU00622.

XX PR 26-MAY-2000; 2000AU-0007761.

XX PA (AUSU ) UNIV AUSTRALIAN NAT.

XX PI Thomson SA, Ramshaw IA;

XX DR WPI; 2002-147575/19.

XX DR N-PSDB; ABK36950.

XX TT New synthetic polypeptides having several different segments of at

XX FT least one parent polypeptide linked together differently compared to

XX FT the linkage in the parent polypeptide, for inducing immune response

XX FT against a pathogen or cancer

XX PS Example 3; Fig 27; 364pp; English.

XX CC The invention relates to a new synthetic polypeptide (I) comprising

XX CC several different segments of at least one parent polypeptide linked

XX CC together in a different relationship relative to their linkage in the

XX CC parent polypeptide to impede, abrogate or otherwise alter at least one

XX CC function associated with the parent polypeptide and for inducing an

XX CC immune response against a pathogen or cancer. Also included are a

XX CC synthetic polynucleotide encoding and a computer system for

XX CC designing the synthetic polypeptides. The synthetic polypeptides and

XX CC polynucleotides are referred to as a Savine. The synthetic polypeptide is

XX CC useful for modulating immune responses preferably directed against a

XX CC pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix,

XX CC colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone

XX CC liver, oesophagus, brain, testicle, uterus), as potentiating agents.

XX CC Compositions comprising the polypeptide may be used in the treatment or

XX CC prophylaxis against viral (such as infections caused by HIV (human

XX CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis

XX CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial

XX CC (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,

XX CC Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic

XX CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,

XX CC trypanosoma, Toxoplasma and Giardia) infections. The present

XX CC sequence is a savine protein of the invention.



QY 1 MAAQERRV 9  
| | | | | | | |  
DB 19 MAAQERRV 27

RESULT 8  
AAU85115  
ID AAU85115 standard; Peptide; 30 AA.

XX AC AAU85115;  
XX DT 08-MAY-2002 (first entry)  
XX DE Human NYSOLB segment 2.  
XX KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;  
KW viral infection; human immunodeficiency virus; melanoma;  
KW bacterial infection; Salmonella; Legionella; parasitic infection;  
KW Trypanosoma; Toxoplasma; Giardia.  
XX OS Homo sapiens.  
XX PN WO200190197-A1.  
XX PD 29-NOV-2001.  
XX PF 25-MAY-2001; 2001WO-AU00622.  
XX PR 26-MAY-2000; 2000AU-0007761.  
XX PA (AUSU ) UNIV AUSTRALIAN NAT.  
XX PI Thomson SA, Ramshaw IA;  
XX DR WPI: 2002-147575/19.  
XX DR N-PSDB; ABK36935.

XX PT New synthetic polypeptides having several different segments of at  
PT least one parent polypeptide linked together differently compared to  
PT the linkage in the parent polypeptide, for inducing immune response  
PT against a pathogen or cancer  
XX PS Example 3; Fig 27; 364pp; English.

XX CC The invention relates to a new synthetic polypeptide (I) comprising  
CC several different segments of at least one parent polypeptide linked  
CC together in a different relationship relative to their linkage in the  
CC parent polypeptide to impede, abrogate or otherwise alter at least one  
CC function associated with the parent polypeptide and for inducing an  
CC immune response against a pathogen or cancer. Also included are a  
CC synthetic polynucleotide encoding and a computer system for  
CC designing the synthetic polypeptides. The synthetic polypeptides and  
CC polynucleotides are referred to as a savine. The synthetic polypeptide is  
CC useful for modulating immune responses preferably directed against a  
CC pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix,  
CC colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone  
CC liver, oesophagus, brain, testicle, uterus), as potentiating agents.  
CC Compositions comprising the polypeptide may be used in the treatment or  
CC prophylaxis against viral (such as infections caused by HIV (human  
CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis  
CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial  
CC (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,  
CC Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic  
CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,  
CC Trypanosoma, Toxoplasma and Giardia) infections. The present  
CC sequence is a peptide derived from a parent protein used to  
CC construct a savine of the invention.

XX SQ Sequence 30 AA;  
Query Match 100.0%; Score 41; DB 23; Length 30;  
Best Local Similarity 100.0%; Pred. No. 0.22; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;

QY 1 MAAQERRV 9  
| | | | | | | |  
DB 4 MAAQERRV 12

RESULT 9  
AAY05966  
ID AAY05966 standard; Protein; 58 AA.

XX AC AAY05966;  
XX DT 16-AUG-1999 (first entry)  
XX DE Human cancer antigen NY ESO-1/CAG-3 ORF2 protein.  
XX KW NY ESO-1/CAG-3 gene; CAG-3 gene; cancer antigen; human; leukaemia;  
KW non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
KW vaccine; ORF2.  
XX OS Homo sapiens.  
XX PN WO9918206-A2.  
XX PF 15-APR-1999.  
XX PR 21-SEP-1998; 98WO-US19609.  
XX PS 08-OCT-1997; 97US-0061428.  
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX PI Rosenberg SA, Wang RF;  
XX DR WPI: 1999-277270/23.  
XX DR N-PSDB; AAX58599.

XX PT Cancer antigen NY ESO1/CAG-3  
XX PS Claim 5; Fig 3A; 88pp; English.

XX CC The present sequence represents the ORF2 protein encoded by  
CC open reading frame 2 of the human ESO-1/CAG-3 (or CAG-3) gene.  
CC CAG-3 is a new and potent tumour antigen capable of eliciting an  
CC antigen specific immune response by T cells. Cancer peptides  
CC comprising ORF2, ORF1 (see AAY05965), portions of these peptides and  
CC their variants (see AAY05967-87), are useful as cancer vaccines that  
CC protect the recipient from development of cancer. The invention  
CC provides: vectors and host cells (also useful as vaccines); a  
CC method of diagnosis of cancer or precancer; a transgenic animal;  
CC antisense oligonucleotides that inhibit expression of the cancer  
CC peptide or tumour antigen; antibodies reacting with the CAG-3  
CC cancer peptide, useful in diagnostic and detection assays; and  
CC methods for preventing or inhibiting cancer by administering a  
CC cancer peptide, with or without an HLA molecule. The cancer  
CC peptides form part of, or are derived from, cancers such as  
CC primary or metastatic melanoma, thymoma, lymphoma, sarcoma, lung  
CC cancer, liver cancer, leukaemia, uterine cancer, cervical cancer,  
CC bladder cancer, kidney cancer and adenocarcinomas such as breast,  
CC prostate, ovarian, pancreatic and thyroid cancers. Melanoma is  
CC treated by inducing cancer-specific T cells in vitro for subsequent  
CC return to a patient.

XX SQ Sequence 58 AA;  
Query Match 100.0%; Score 41; DB 20; Length 58;  
Best Local Similarity 100.0%; Pred. No. 0.42; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;

```
XX SQ Sequence 13 AA;
Query Match 100.0%; Score 41; DB 20; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.094;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAQERRV 9
Db 3 MAAQERRV 11
|||||

RESULT 6
AAY70904
ID AAY70904 standard; peptide; 19 AA.
XX AC AAY70904;
XX DT 31-JUL-2000 (first entry)
XX OS Homo sapiens.
XX OS Synthetic.
XX PN WO200023584-A1.
XX PD 27-APR-2000.
XX PF 15-OCT-1999; 99WO-EP07832.
XX PR 16-OCT-1998; 98EP-0119583.
XX PA (BOEH ) BOEHRINGER INGELHEIM INT GMBH.
XX PA (UYHO-) UNIV HOSPITAL LEIDEN.
XX PI Schrier PI, Aarnoudse CA, Heider K, Klade C;
XX WPI: 2000-339685/29.

Tumor-associated antigen useful for cancer immunotherapy is encoded by
the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA -.
Disclosure; Page 17; 73pp; English.

The present sequence is the synthetic peptide F4, derived from the
hydrophobic regions of human CAMEL (Cytotoxic T lymphocytes (CTL)-
recognised Antigen on MELanoma), a tumour-associated antigen. It
corresponds to residues 15-31 of the CAMEL protein and is used to raise
antibodies. The Lys residue at both termini enable the peptides to be
linked to the carrier keyhole limpet haemocyanin (KLH) for immunisation.
CAMEL protein is encoded by the LAGE-1 gene, a tumour-specific antigen.
It shows strong homology with NY-ESO-1, a melanoma specific tumour
antigen. The tumour-associated antigen displayed on melanoma cells is
recognised by cytotoxic T lymphocytes. CAMEL is expressed in tumour cell
lines, tumour tissues (e.g. breast and lung) and in restricted number of
healthy tissues. This sequence has anticancer activity. CAMEL tumour
antigen and immunogenic peptides derived from it are useful for cancer
immunotherapy. They have the potential to induce an immune response, by
eliciting a CTL response. The DNA molecule is used to construct
recombinant or fusion proteins.

XX SQ Sequence 19 AA;
Query Match 100.0%; Score 41; DB 21; Length 19;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAQERRV 9
```

```
Db |||||
4 MAAQERRV 12
```

RESULT 7

AAU85114

ID AAU85114 standard; Peptide; 30 AA.

XX AC AAU85114;

XX DT 08-MAY-2002 (first entry)

XX DE Human NYNSO1b segment 1.

XX KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;

XX KW viral infection; human immunodeficiency virus; melanoma;

XX KW bacterial infection; Salmonella; Legionella; parasitic infection;

XX KW Trypanosoma; Toxoplasma; Giardia.

XX OS Homo sapiens.

XX PN WO200190197-A1.

XX PD 29-NOV-2001.

XX PF 25-MAY-2001; 2001WO-AU00622.

XX PR 26-MAY-2000; 2000AU-0007761.

XX PA (AUSU ) UNIV AUSTRALIAN NAT.

XX PI Thomson SA, Ramshaw IA;

XX WPI: 2002-147575/19.

XX N-PSDB; ABK36934.

XX New synthetic polypeptides having several different segments of at

least one parent polypeptide linked together differently compared to

the linkage in the parent polypeptide, for inducing immune response

against a pathogen or cancer

Example 3; Fig 27; 364pp; English.

The invention relates to a new synthetic polypeptide (I) comprising

several different segments of at least one parent polypeptide linked

together in a different relationship relative to their linkage in the

parent polypeptide to impede, abrogate or otherwise alter at least one

function associated with the parent polypeptide and for inducing an

immune response against a pathogen or cancer. Also included are a

synthetic polynucleotide encoding and a computer system for

designing the synthetic polypeptides. The synthetic polypeptides and

polynucleotides are referred to as a vaccine. The synthetic polypeptide is

useful for modulating immune responses preferably directed against a

pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix,

colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone

liver, oesophagus, brain, testicle, uterus), as potentiating agents.

Compositions comprising the polypeptide may be used in the treatment or

prophylaxis against viral (such as infections caused by HIV (human

immunodeficiency virus), hepatitis, influenza, Japanese encephalitis

virus, Epstein-Barr virus and respiratory syncytial virus), bacterial

(e.g., infections caused by Neisseria, Meningococcus, Haemophilus,

Salmonella, Streptococcus, Legionella and Mycobacterium or parasitic

(e.g., infections caused by Plasmodium, Schistosoma, Leishmania,

Trypanosoma, Toxoplasma and Giardia) infections. The present

sequence is a peptide derived from a parent protein used to

construct a vaccine of the invention.

XX SQ Sequence 30 AA;

Query Match 100.0%; Score 41; DB 23; Length 30;

Best Local Similarity 100.0%; Pred. No. 0.22; Mismatches 0; Indels 0; Gaps 0;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC from, cancers such as primary or metastatic melanoma, thymoma, lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine cancer, cervical cancer, bladder cancer, kidney cancer and adenocarcinomas such as breast, prostate, ovarian, pancreatic and thyroid cancers. Melanoma is treated by inducing cancer-specific T cells in vitro for subsequent return to a patient.

CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
SQ Sequence 11 AA; Query Match 100.0%; Score 41; DB 20; Length 11; Best Local Similarity 100.0%; Pred. No. 0.08; Mismatches 0; Indels 0; Gaps 0; Matches 9; Conservative 0

QY 1 MAAQERRV 9  
Db 1 MAAQERRV 9  
IIIIIIII

RESULT 4  
AAY05984  
ID AAY05984 standard; Peptide; 12 AA.  
AC AAY05984;  
XX  
XX 16-AUG-1999 (first entry)  
DT Human cancer antigen NY ESO-1/CAG-3 ORF2 cancer peptide.  
DE  
XX NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human; leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer; metastasis; melanoma; adenocarcinoma; thymoma; colon cancer; uterine cancer; breast cancer; prostate cancer; ovarian cancer; cervical cancer; bladder cancer; kidney cancer; pancreatic cancer; liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy; vaccine.  
XX Homo sapiens.  
XX WO9918206-A2.  
XX 15-APR-1999.  
XX 21-SEP-1998; 98WO-US19609.  
XX 08-OCT-1997; 97US-0061428.  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX Rosenberg SA, Wang RF;  
XX WPI; 1999-277270/23.  
XX Cancer antigen NY ESO1/CAG-3  
XX Disclosure; Page 12; 88pp; English.

CC The present sequence represents a cancer peptide that corresponds to amino acid residues 16-27 of human ESO-1/CAG-3 (or CAG-3) ORF2 (see AAY05966), a new and potent tumour antigen capable of eliciting an antigen specific immune response by T cells. Cancer peptides derived from CAG-3 ORF2, CAG-3 ORF1 (see AAY05965), portions of them and their variants (see AAY05967-87), are useful as cancer vaccines that protect against cancer. The invention provides: vectors and host cells (also useful as vaccines); a method of diagnosis of cancer or precancer; a transgenic animal; antisense oligonucleotides that inhibit expression of the cancer peptide or tumour antigen; antibodies reacting with a CAG-3 cancer peptide, useful in diagnostic and detection assays; and methods for preventing or inhibiting cancer by administering a cancer peptide, with or without an HLA molecule. The cancer peptides form part of, or are derived from, cancers such as primary or metastatic melanoma, thymoma, lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine cancer, cervical cancer, bladder cancer, kidney cancer and adenocarcinomas such as breast, prostate, ovarian, pancreatic and thyroid cancers. Melanoma is treated by inducing cancer-specific T cells in vitro for subsequent return to a patient.

CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and thyroid cancers. Melanoma is treated by inducing cancer-specific T cells in vitro for subsequent return to a patient.

CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
SQ Sequence 12 AA; Query Match 100.0%; Score 41; DB 20; Length 12; Best Local Similarity 100.0%; Pred. No. 0.087; Mismatches 0; Indels 0; Gaps 0; Matches 9; Conservative 0

QY 1 MAAQERRV 9  
Db 2 MAAQERRV 10  
IIIIIIII

RESULT 5  
AAY05985  
ID AAY05985 standard; Peptide; 13 AA.  
XX AAY05985;  
AC  
XX 16-AUG-1999 (first entry)  
DT Human cancer antigen NY ESO-1/CAG-3 ORF2 cancer peptide.  
DE  
XX NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human; leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer; metastasis; melanoma; adenocarcinoma; thymoma; colon cancer; uterine cancer; breast cancer; prostate cancer; ovarian cancer; cervical cancer; bladder cancer; kidney cancer; pancreatic cancer; liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy; vaccine.  
XX Homo sapiens.  
XX WO9918206-A2.  
XX 15-APR-1999.  
XX 21-SEP-1998; 98WO-US19609.  
XX 08-OCT-1997; 97US-0061428.  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX Rosenberg SA, Wang RF;  
XX WPI; 1999-277270/23.  
XX Cancer antigen NY ESO1/CAG-3  
XX Disclosure; Page 12; 88pp; English.

CC The present sequence represents a cancer peptide that corresponds to amino acid residues 15-27 of human ESO-1/CAG-3 (or CAG-3) ORF2 (see AAY05966), a new and potent tumour antigen capable of eliciting an antigen specific immune response by T cells. Cancer peptides derived from CAG-3 ORF2, CAG-3 ORF1 (see AAY05965), portions of them and their variants (see AAY05967-87), are useful as cancer vaccines that protect against cancer. The invention provides: vectors and host cells (also useful as vaccines); a method of diagnosis of cancer or precancer; a transgenic animal; antisense oligonucleotides that inhibit expression of the cancer peptide or tumour antigen; antibodies reacting with a CAG-3 cancer peptide, useful in diagnostic and detection assays; and methods for preventing or inhibiting cancer by administering a cancer peptide, with or without an HLA molecule. The cancer peptides form part of, or are derived from, cancers such as primary or metastatic melanoma, thymoma, lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine cancer, cervical cancer, bladder cancer, kidney cancer and adenocarcinomas such as breast, prostate, ovarian, pancreatic and thyroid cancers. Melanoma is treated by inducing cancer-specific T cells in vitro for subsequent return to a patient.

PT the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA -  
 PS Claim 8; Page 34; 73pp; English.  
 CC The present sequence is an immunogenic peptide CAMEL 17, of the human  
 CC tumour-associated antigen CAMEL (Cytotoxic T lymphocytes (CTL)-recognised  
 CC Antigen on MELanoma). This peptide has the potential to bind to HLA-A2  
 CC and corresponds to residues 17-25 of the CAMEL protein. The CAMEL protein  
 CC is encoded by the LAGE-1 gene, a tumour-specific antigen. It is different  
 CC from the LAGE-1 protein, since it is translated from a different open  
 CC reading frame (ORF-1). It shows strong homology with NY-ESO-1, a melanoma  
 CC specific tumour antigen. The tumour-associated antigen displayed on  
 CC melanoma cells is recognised by cytotoxic T lymphocytes. CAMEL is  
 CC expressed in tumour cell lines, tumour tissues (e.g. breast and lung) and  
 CC in restricted number of healthy tissues. This sequence has anticancer  
 CC activity. CAMEL tumour antigen and immunogenic peptides derived from it  
 CC are useful for cancer immunotherapy. They have the potential to induce an  
 CC immune response, by eliciting a CTL response. The DNA molecule is used to  
 CC construct recombinant or fusion proteins.  
 CC  
 SQ Sequence 9 AA;  
 Query Match 100.0%; Score 41; DB 21; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MLAAQERRV 9  
 Db 1 MLAAQERRV 9  
 DE CAMEL16 immunogenic peptide of human CAMEL protein.  
 DE CAMEL: CTL-recognised Antigen on MELanoma; cytotoxic T lymphocyte; CTL;  
 KW tumour-associated antigen; LAGE-1; NY-ESO-1; anticancer; melanoma;  
 KW human; cancer; immunotherapy; immunogenic peptide; immune response.  
 XX Homo sapiens.  
 XX WO200023584-A1.  
 XX 27-APR-2000.  
 XX 15-OCT-1999; 99WO-EP07832.  
 XX 16-OCT-1998; 98EP-0119583.  
 XX (BOEH) BOEHRINGER INGELHEIM INT GMBH.  
 PA (UYHO-) UNIV HOSPITAL LEIDEN.  
 XX Schrier PI, Aarnoudse CA, Heider K, Klade C;  
 XX WPI; 2000-339685/29.  
 XX Tumour-associated antigen useful for cancer immunotherapy is encoded by  
 PT the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA -  
 PS Claim 7; Page 34; 73pp; English.  
 CC The present sequence is an immunogenic peptide CAMEL 16, of the human  
 CC tumour-associated antigen CAMEL (Cytotoxic T lymphocytes (CTL)-recognised  
 CC Antigen on MELanoma). This peptide has the potential to bind to HLA-A2  
 CC and corresponds to residues 16-25 of the CAMEL protein. The CAMEL protein  
 CC is encoded by the LAGE-1 gene, a tumour-specific antigen. It is different  
 CC from the LAGE-1 protein, since it is translated from a different open

CC reading frame (ORF-1). It shows strong homology with NY-ESO-1, a melanoma  
 CC specific tumour antigen. The tumour-associated antigen displayed on  
 CC melanoma cells is recognised by cytotoxic T lymphocytes. CAMEL is  
 CC expressed in tumour cell lines, tumour tissues (e.g. breast and lung) and  
 CC in restricted number of healthy tissues. This sequence has anticancer  
 CC activity. CAMEL tumour antigen and immunogenic peptides derived from it  
 CC are useful for cancer immunotherapy. They have the potential to induce an  
 CC immune response, by eliciting a CTL response. The DNA molecule is used to  
 CC construct recombinant or fusion proteins.  
 CC  
 SQ Sequence 10 AA;  
 Query Match 100.0%; Score 41; DB 21; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.072;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MLAAQERRV 9  
 Db 2 MLAAQERRV 10  
 DE Human cancer antigen NY ESO-1/CAG-3 ORF2 cancer peptide.  
 DE NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;  
 KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
 KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
 KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
 KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
 KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
 XX vaccine.  
 XX Homo sapiens.  
 XX WO9918206-A2.  
 XX 15-APR-1999.  
 XX 21-SEP-1998; 98WO-US19609.  
 XX 08-OCT-1997; 97US-0061428.  
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.  
 PA Rosenberg SA, Wang RF;  
 XX WPI; 1999-277270/23.  
 XX Cancer antigen NY ESO1/CAG-3  
 XX Disclosure; Page 12; 88pp; English.  
 CC The present sequence represents a cancer peptide that corresponds  
 CC to amino acid residues 17-27 of human ESO-1/CAG-3 (or CAG-3) ORF2  
 CC (see AAY05966), a new and potent tumour antigen capable of eliciting  
 CC an antigen specific immune response by T cells. Cancer peptides  
 CC derived from CAG-3 ORF2, CAG-3 ORF1 (see AAY05965), portions of them  
 CC and their variants (see AAY05967-87), are useful as cancer vaccines  
 CC that protect against cancer. The invention provides: vectors and  
 CC host cells (also useful as vaccines); a method of diagnosis of  
 CC cancer or precancer; a transgenic animal; antisense oligonucleotides  
 CC that inhibit expression of the cancer peptide or tumour antigen;  
 CC antibodies reacting with a CAG-3 cancer peptide, useful in  
 CC diagnostic and detection assays; and methods for preventing or  
 CC inhibiting cancer by administering a cancer peptide, with or without  
 CC an HLA molecule. The cancer peptides form part of, or are derived

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:04:57 ; Search time 14.9241 Seconds  
(without alignments)  
95.721 Million cell updates/sec

Title: US-09-807-512-26  
Perfect score: 41  
Sequence: 1 MAAQERRV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues 1107863  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A_Geneseq_19Jun03.*			
1:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*		
2:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*		
3:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*		
4:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*		
5:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*		
6:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*		
7:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*		
8:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*		
9:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*		
10:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*		
11:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*		
12:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*		
13:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*		
14:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*		
15:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*		
16:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*		
17:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*		
18:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*		
19:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*		
20:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*		
21:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*		
22:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*		
23:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*		
24:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	100.0	9	AAV70859	CAMEL17 immunogeni
2	41	100.0	10	AAV70858	CAMEL16 immunogeni
3	41	100.0	11	AAV05983	Human cancer anti
4	41	100.0	12	AAV05984	Human cancer anti
5	41	100.0	13	AAV05985	Human cancer anti
6	41	100.0	19	AAV70904	Synthetic peptide,
7	41	100.0	30	AAU85114	Human NYNSO1b segm
8	41	100.0	30	AAU85115	Human NYNSO1b segm
9	41	100.0	58	AAV05966	Human cancer anti

10	41	100.0	58	21	AAV70863	Human tumour anti
11	41	100.0	58	23	AAU84819	Human NYNSO1b cons
12	41	100.0	109	21	AAV70854	Human CTL-recognis
13	41	100.0	3541	23	AAU85130	Human melanoma spe
14	36	87.8	10	20	AAV05982	Human cancer anti
15	36	87.8	10	22	AAV05982	Exemplary antigen
16	36	87.8	10	23	ABG6803	Tumour antigen NY-
17	32	78.0	9	20	AAV05981	Human cancer anti
18	32	78.0	198	23	ABP41770	Human ovarian anti
19	32	78.0	456	22	AAU84081	Human protein sequ
20	32	78.0	673	22	AAV37106	Amino acid sequenc
21	31	75.6	579	22	ABB60380	Drosophila melanog
22	30	73.2	661	23	ABB09587	Rat HRC-1 protein.
23	30	73.2	758	22	ABG6803	mMfn2. Mus muscul
24	30	73.2	758	24	ABU64938	Mouse mitofusin 2,
25	29	70.7	80	20	AAV11476	Human 5' EST secre
26	29	70.7	127	22	ABG18982	Novel human diagno
27	29	70.7	141	22	AAV72604	Human Electron Tra
28	29	70.7	153	22	ABB96061	Human testicular a
29	29	70.7	153	22	AAV95369	Human reproductive
30	29	70.7	161	22	ABG15245	Novel human diagno
31	29	70.7	187	22	ABG02389	Novel human diagno
32	29	70.7	196	23	ABV97861	Human secretory po
33	29	70.7	204	23	ABU05910	M. tuberculosis an
34	29	70.7	226	24	ABU00156	Human novel polype
35	29	70.7	263	18	AAW18783	Cytohesin 2 encode
36	29	70.7	267	24	ABR48176	Human bladder can
37	29	70.7	270	20	AAV34117	Soybean partial cy
38	29	70.7	271	22	ABG18979	Novel human diagno
39	29	70.7	282	22	AAU34080	Staphylococcus aur
40	29	70.7	283	22	AAU52245	Propionibacterium
41	29	70.7	292	22	AAU36847	Staphylococcus aur
42	29	70.7	333	15	AAU48691	G-protein coupled
43	29	70.7	333	17	AAW02663	G-protein coupled
44	29	70.7	355	23	AAE15535	Beta vulgaris RNA
45	29	70.7	379	22	ABB70869	Drosophila melanog

ALIGNMENTS

RESULT 1	
AAV70859	AAV70859 standard; peptide; 9 AA.
ID	AAV70859 standard; peptide; 9 AA.
XX	
AC	AAV70859;
XX	
DT	31-JUL-2000 (first entry)
XX	
DE	CAMEL17 immunogenic peptide of human CAMEL protein.
XX	
KW	CAMEL; CTL-recognised Antigen on MELANOMA; cytotoxic T lymphocyte; CTL;
KW	tumour-associated antigen; LAGE-1; NY-ESO-1; anticancer; melanoma;
KW	human; cancer; immunotherapy; immunogenic peptide; immune response.
XX	
OS	Homo sapiens.
XX	
PN	WO200023584-A1.
XX	
PD	27-APR-2000.
XX	
PF	15-OCT-1999; 99WO-EF07832.
XX	
PR	16-OCT-1998; 98EP-0119583.
XX	
PA	(BOEH ) BOEHRINGER INGELHEIM INT GMBH.
PA	(UYHO-) UNIV HOSPITAL LEIDEN.
XX	
PI	Schrier PI, Aarnoudse CA, Heider K, Klade C;
XX	
DR	WPI; 2000-339685/29.
XX	
PT	Tumor-associated antigen useful for cancer immunotherapy is encoded by



**THIS PAGE BLANK (USPTO)**

```
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8415
; SEQ ID NO 5897
; LENGTH: 2603
; TYPE: PRT
; ORGANISM: Xenorhabdus sp.
US-09-897-516A-5897

Query Match      71.1%; Score 32; DB 5; Length 2603;
Best Local Similarity 60.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 AMLAQOERV 10
Db      617 ALIADDDRV 626

Search completed: October 7, 2003, 13:55:57
Job time : 4.16456 secs
```

; PRIOR APPLICATION NUMBER: 60/406,640  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,588  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,576  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,646  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,666  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,653  
; PRIOR FILING DATE: 2002-08-29  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 3700  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2438  
; LENGTH: 794  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US03-26780-2438

Query Match 73.3%; Score 33; DB 1; Length 794;  
Best Local Similarity 70.0%; Pred. No. 67;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AMLAAQERRV 10  
| : : : :  
Db 300 AFLASKERRV 309

RESULT 12  
PCT-US03-26780-2245  
; Sequence 2245, Application PC/TUS0326780  
; GENERAL INFORMATION:  
; APPLICANT: FIVEPRIME THERAPEUTICS, INC.  
; TITLE OF INVENTION: HUMAN POLYPEPTIDES ENCODED BY POLYNUCLEOTIDES AND METHODS OF  
; FILE OF INVENTION: THEIR USE  
; FILE REFERENCE: 08940.0014-00304  
; CURRENT APPLICATION NUMBER: PCT/US03/26780  
; CURRENT FILING DATE: 2003-08-28  
; PRIOR APPLICATION NUMBER: 60/406,616  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,579  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,655  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,642  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,640  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,588  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,576  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,646  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,666  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,653  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 3700  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2245  
; LENGTH: 354  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US03-26780-2245

Query Match 71.1%; Score 32; DB 1; Length 354;  
Best Local Similarity 77.8%; Pred. No. 54;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AMLAAQERR 9  
| : : : :  
Db 323 AMLAAQRRQ 331

RESULT 13  
US-10-425-114A-70170  
; Sequence 70170, Application US/10425114A  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114A  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 70170  
; LENGTH: 736  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-2MFLE73120C10\_FLI.pep  
US-10-425-114A-70170

Query Match 71.1%; Score 32; DB 6; Length 736;  
Best Local Similarity 100.0%; Pred. No. 94;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 AAQERRV 10  
| : : : :  
Db 568 AAQERRV 574

RESULT 14  
US-09-786-108A-15  
; Sequence 15, Application US/09786108A  
; GENERAL INFORMATION:  
; APPLICANT: DIXON, JANE  
; APPLICANT: MCKINNON, DAVID  
; TITLE OF INVENTION: Mammalian elk Potassium Channel Genes  
; FILE REFERENCE: 08874138PC1  
; CURRENT APPLICATION NUMBER: US/09/786,108A  
; CURRENT FILING DATE: 2003-05-06  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 1163  
; TYPE: PRT  
; ORGANISM: RAT  
; FEATURE:  
; OTHER INFORMATION: erg 1  
US-09-786-108A-15

Query Match 71.1%; Score 32; DB 5; Length 1163;  
Best Local Similarity 60.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AMLAAQERRV 10  
| : : : :  
Db 85 ALLGAERKV 94

RESULT 15  
US-09-897-516A-5897  
; Sequence 5897, Application US/09897516A  
; GENERAL INFORMATION:  
; APPLICANT: Corbin, David R.  
; APPLICANT: Goldman, Barry S.

```

; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 68789
; LENGTH: 450
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700342573_FLI.pep
US-10-425-114A-68789

Query Match 75.6%; Score 34; DB 6; Length 450;
Best Local Similarity 77.8%; Pred. No. 29;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 AMLAAQERR 9
DB 432 AMLASQKR 440

RESULT 8
US-10-425-114A-46933
; Sequence 46933, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 46933
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700620710_FLI.pep
US-10-425-114A-46933

Query Match 73.3%; Score 33; DB 6; Length 162;
Best Local Similarity 80.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AMLAAQERR 10
DB 146 AELEAQERR 155

RESULT 9
US-10-425-114A-67943
; Sequence 67943, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
```

```

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 67943
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-2MFLM017186F05_FLI.pep
US-10-425-114A-67943

Query Match 73.3%; Score 33; DB 6; Length 164;
Best Local Similarity 80.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AMLAAQERR 10
DB 148 AELEAQERR 157

RESULT 10
US-10-425-114A-64836
; Sequence 64836, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 64836
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4720-012-D11_FLI.pep
US-10-425-114A-64836

Query Match 73.3%; Score 33; DB 6; Length 446;
Best Local Similarity 60.0%; Pred. No. 43;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 AMLAAQERR 10
DB 64 SLLAAQERR 73

RESULT 11
PCT-US03-26780-2438
; Sequence 2438, Application PC/TUS0326780
; GENERAL INFORMATION:
; APPLICANT: FIVEPRIME THERAPEUTICS, INC.
; TITLE OF INVENTION: HUMAN POLYPEPTIDES ENCODED BY POLYNUCLEOTIDES AND METHODS OF
; FILE REFERENCE: 08940.0014-00304
; CURRENT APPLICATION NUMBER: PCT/US03/26780
; CURRENT FILING DATE: 2003-08-28
; PRIOR APPLICATION NUMBER: 60/406,616
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/406,579
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/406,655
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/406,642
; PRIOR FILING DATE: 2002-08-29
```

```
; FEATURE:
; OTHER INFORMATION: NYSOLB segment 2
US-10-296-734-1424

Query Match      100.0%; Score 45; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.043; 0; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0;

QY 1 AMLAAQERRV 10
Db 3 AMLAAQERRV 12

RESULT 3
US-10-296-734-833
; Sequence 833, Application US/10296734
; GENERAL INFORMATION:
; APPLICANT: Thompson, Scott A
; APPLICANT: Ramshaw, Ian A
; TITLE OF INVENTION: Synthetic molecules and uses therefor
; FILE REFERENCE: Savine
; CURRENT APPLICATION NUMBER: US/10/296,734
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: AU PQ7761/00
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 833
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: NYSOLB consensus polypeptide
US-10-296-734-833

Query Match      100.0%; Score 45; DB 6; Length 58;
Best Local Similarity 100.0%; Pred. No. 0.07; 0; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0;

QY 1 AMLAAQERRV 10
Db 16 AMLAAQERRV 25

RESULT 4
US-10-296-734-1454
; Sequence 1454, Application US/10296734
; GENERAL INFORMATION:
; APPLICANT: Thompson, Scott A
; APPLICANT: Ramshaw, Ian A
; TITLE OF INVENTION: Synthetic molecules and uses therefor
; FILE REFERENCE: Savine
; CURRENT APPLICATION NUMBER: US/10/296,734
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: AU PQ7761/00
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1454
; LENGTH: 3541
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Melanoma cancer specific savine
US-10-296-734-1454

Query Match      100.0%; Score 45; DB 6; Length 3541;
Best Local Similarity 100.0%; Pred. No. 1.6; 0; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0;

QY 1 AMLAAQERRV 10
Db 183 AMLAAQERRV 192

RESULT 5
US-10-654-102-35
; Sequence 35, Application US/10654102
; GENERAL INFORMATION:
; APPLICANT: CHAN, LAWRENCE
; APPLICANT: KOJIMA, HIDEOTO
; TITLE OF INVENTION: INDUCTION OF PANCREATIC ISLET FORMATION
; FILE REFERENCE: P02409US1
; CURRENT APPLICATION NUMBER: US/10/654,102
; CURRENT FILING DATE: 2003-09-03
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (22)..(94)
; OTHER INFORMATION: x = anything
US-10-654-102-35

Query Match      75.6%; Score 34; DB 6; Length 103;
Best Local Similarity 80.0%; Pred. No. 9.4; 2; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 2;

QY 1 AMLAAQERRV 10
Db 28 AAXAAQERRV 37

RESULT 6
US-10-425-114A-71411
; Sequence 71411, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 71411
; LENGTH: 317
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73091F09_FLI.pep
US-10-425-114A-71411

Query Match      75.6%; Score 34; DB 6; Length 317;
Best Local Similarity 77.8%; Pred. No. 22; 2; Indels 0; Gaps 0;
Matches 7; Conservative 2; Mismatches 0;

QY 1 AMLAAQERR 9
Db 297 AMLAQRRR 305

RESULT 7
US-10-425-114A-68789
; Sequence 68789, Application US/10425114A
; GENERAL INFORMATION:
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:26:33 ; Search time 3.16456 Seconds  
(without alignments)  
92.039 Million cell updates/sec

Title: US-09-807-512-25

Perfect score: 45

Sequence: 1 AMLAAQERRV 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 136899 seqs, 29126274 residues

Total number of hits satisfying chosen parameters: 136899

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_New:\*

- 1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	45	100.0	30	6	US-10-296-734-1422 Sequence 1422, Ap
2	45	100.0	30	6	US-10-296-734-1424 Sequence 1424, Ap
3	45	100.0	58	6	US-10-296-734-833 Sequence 833, App
4	45	100.0	3541	6	US-10-296-734-1454 Sequence 1454, Ap
5	34	75.6	103	6	US-10-654-102-35 Sequence 35, Appl
6	34	75.6	317	6	US-10-425-114A-71411 Sequence 71411, A
7	34	75.6	450	6	US-10-425-114A-68789 Sequence 68789, A
8	33	73.3	162	6	US-10-425-114A-46933 Sequence 46933, A
9	33	73.3	164	6	US-10-425-114A-67943 Sequence 67943, A
10	33	73.3	446	6	US-10-425-114A-64836 Sequence 64836, A
11	33	73.3	794	1	PCT-US03-26780-2438 Sequence 2438, Ap
12	32	71.1	354	1	PCT-US03-26780-2245 Sequence 2245, Ap
13	32	71.1	736	6	US-10-425-114A-70170 Sequence 70170, A
14	32	71.1	1163	5	US-09-786-108A-15 Sequence 15, Appl
15	32	71.1	2603	5	US-09-897-516A-5897 Sequence 5897, Ap
16	31	68.9	347	1	PCT-US03-24669-286 Sequence 286, App
17	31	68.9	545	1	PCT-US03-18787-135 Sequence 135, App
18	31	68.9	545	1	PCT-US03-19089-135 Sequence 135, App
19	30	66.7	545	5	US-09-897-516A-4422 Sequence 4422, Ap
20	30	66.7	1125	5	US-09-897-516A-4515 Sequence 4515, Ap
21	29	64.4	267	7	US-60-490-890-2340 Sequence 2340, Ap
22	29	64.4	274	6	US-10-425-114A-49793 Sequence 49793, A
23	29	64.4	350	6	US-10-296-115-877 Sequence 877, App
24	29	64.4	407	6	US-10-425-114A-70887 Sequence 70887, A
25	29	64.4	492	6	US-10-425-114A-65735 Sequence 65735, A
26	29	64.4	504	6	US-10-425-114A-70248 Sequence 70248, A

#### ALIGNMENTS

##### RESULT 1

US-10-296-734-1422  
; Sequence 1422, Application US/10296734  
; GENERAL INFORMATION:  
; APPLICANT: Thompson, Scott A  
; TITLE OF INVENTION: Synthetic molecules and uses therefor  
; FILE REFERENCE: Savine  
; CURRENT APPLICATION NUMBER: US/10/296,734  
; CURRENT FILING DATE: 2003-08-04  
; PRIOR APPLICATION NUMBER: AU PQ7761/00  
; PRIOR FILING DATE: 2000-05-26  
; NUMBER OF SEQ ID NOS: 1507  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1422  
; LENGTH: 30  
; TYPE: PRT  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: NYSolb segment 1  
US-10-296-734-1422

Query Match 100.0%; Score 45; DB 6; Length 30;  
Best Local Similarity 100.0%; Pred. No. 0.043; 0; Indels 0; Gaps 0;  
Matches 10; Conservative 0; Mismatches 0;

Qy 1 AMLAAQERRV 10  
|||||  
Db 18 AMLAAQERRV 27

##### RESULT 2

US-10-296-734-1424  
; Sequence 1424, Application US/10296734  
; GENERAL INFORMATION:  
; APPLICANT: Thompson, Scott A  
; TITLE OF INVENTION: Synthetic molecules and uses therefor  
; FILE REFERENCE: Savine  
; CURRENT APPLICATION NUMBER: US/10/296,734  
; CURRENT FILING DATE: 2003-08-04  
; PRIOR APPLICATION NUMBER: AU PQ7761/00  
; PRIOR FILING DATE: 2000-05-26  
; NUMBER OF SEQ ID NOS: 1507  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1424  
; LENGTH: 30  
; TYPE: PRT  
; ORGANISM: Artificial

**THIS PAGE BLANK (USPTO)**

Best Local Similarity 100.0%; Pred. No. 0.32;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AMLAAQERRV 10  
          |||||  
Db 16 AMLAAQERRV 25

Search completed: October 7, 2003, 13:55:01  
Job time : 79.7342 secs



; GENERAL INFORMATION:  
; APPLICANT: Wang Rong, Fu  
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene  
; FILE REFERENCE: 20264269US1  
; CURRENT APPLICATION NUMBER: US/09/529,206A  
; CURRENT FILING DATE: 2000-06-13  
; PRIOR APPLICATION NUMBER: PCT/US98/19609  
; PRIOR FILING DATE: 1998-09-21  
; PRIOR APPLICATION NUMBER: US60/061,428  
; NUMBER OF SEQ ID NOS: 127  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 56  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-206A-56

Query Match 100.0%; Score 45; DB 19; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AMLAAQERRV 10  
Db 16 AMLAAQERRV 25

RESULT 12  
US-09-529-206B-56  
; Sequence 56, Application US/09529206B  
; GENERAL INFORMATION:  
; APPLICANT: Wang Rong, Fu  
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene  
; FILE REFERENCE: 20264269US1  
; CURRENT APPLICATION NUMBER: US/09/529,206B  
; CURRENT FILING DATE: 2000-06-13  
; PRIOR APPLICATION NUMBER: PCT/US98/19609  
; PRIOR FILING DATE: 1998-09-21  
; PRIOR APPLICATION NUMBER: US60/061,428  
; NUMBER OF SEQ ID NOS: 127  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 56  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-206B-56

Query Match 100.0%; Score 45; DB 19; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AMLAAQERRV 10  
Db 16 AMLAAQERRV 25

RESULT 13  
US-09-529-206D-56  
; Sequence 56, Application US/09529206D  
; GENERAL INFORMATION:  
; APPLICANT: Wang Rong, Fu  
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene  
; FILE REFERENCE: 218791  
; CURRENT APPLICATION NUMBER: US/09/529,206D  
; CURRENT FILING DATE: 2000-06-13  
; PRIOR APPLICATION NUMBER: PCT/US98/19609

; PRIOR FILING DATE: 1998-09-21  
; PRIOR APPLICATION NUMBER: US60/061,428  
; PRIOR FILING DATE: 1997-10-08  
; NUMBER OF SEQ ID NOS: 127  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 56  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-206D-56

Query Match 100.0%; Score 45; DB 19; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AMLAAQERRV 10  
Db 16 AMLAAQERRV 25

RESULT 14  
US-09-529-206-5  
; Sequence 5, Application US/09529206  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America as  
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene  
; FILE REFERENCE: 20264269PC  
; CURRENT APPLICATION NUMBER: US/09/529,206  
; CURRENT FILING DATE: 2000-06-13  
; PRIOR APPLICATION NUMBER: US60/061,428  
; PRIOR FILING DATE: 1997-10-08  
; NUMBER OF SEQ ID NOS: 106  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 58  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-206-5

Query Match 100.0%; Score 45; DB 19; Length 58;  
Best Local Similarity 100.0%; Pred. No. 0.32;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AMLAAQERRV 10  
Db 16 AMLAAQERRV 25

RESULT 15  
US-09-529-206A-5  
; Sequence 5, Application US/09529206A  
; GENERAL INFORMATION:  
; APPLICANT: Wang Rong, Fu  
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene  
; FILE REFERENCE: 20264269US1  
; CURRENT APPLICATION NUMBER: US/09/529,206A  
; CURRENT FILING DATE: 2000-06-13  
; PRIOR APPLICATION NUMBER: PCT/US98/19609  
; PRIOR FILING DATE: 1998-09-21  
; PRIOR APPLICATION NUMBER: US60/061,428  
; NUMBER OF SEQ ID NOS: 127  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 58  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-206A-5

Query Match 100.0%; Score 45; DB 19; Length 58;

```

; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206-50

Query Match      100.0%; Score 45; DB 19; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.055;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AMLAAQERRV 10
        |||||
Db      2 AMLAAQERRV 11

RESULT 7
US-09-529-206A-50
; Sequence 50, Application US/09529206A
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; FILE REFERENCE: 20264269US1
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: PCT/US98/19609
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 50
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206A-50

Query Match      100.0%; Score 45; DB 19; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.055;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AMLAAQERRV 10
        |||||
Db      2 AMLAAQERRV 11

RESULT 8
US-09-529-206B-50
; Sequence 50, Application US/09529206B
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; FILE REFERENCE: 20264269US1
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: PCT/US98/19609
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 50
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206B-50

Query Match      100.0%; Score 45; DB 19; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.055;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AMLAAQERRV 10
        |||||
Db      2 AMLAAQERRV 11

RESULT 9
US-09-529-206D-50
; Sequence 50, Application US/09529206D
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; FILE REFERENCE: 218791
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: PCT/US98/19609
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 50
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206D-50

Query Match      100.0%; Score 45; DB 19; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.055;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AMLAAQERRV 10
        |||||
Db      2 AMLAAQERRV 11

RESULT 10
US-09-529-206-56
; Sequence 56, Application US/09529206
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America as
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; FILE REFERENCE: 20264269PC
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 56
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206-56

Query Match      100.0%; Score 45; DB 19; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.12; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AMLAAQERRV 10
        |||||
Db      16 AMLAAQERRV 25

RESULT 11
US-09-529-206A-56
; Sequence 56, Application US/09529206A
```

```
Query Match      100.0%; Score 45; DB 23; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AMLAAQERV 10
   |||||
Db 1 AMLAAQERV 10

RESULT 2
US-09-529-206-49
; Sequence 49, Application US/09529206
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; FILE REFERENCE: 20264269US1
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US/09/529,206
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 49
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206-49

Query Match      100.0%; Score 45; DB 19; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.05;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AMLAAQERV 10
   |||||
Db 1 AMLAAQERV 10

RESULT 3
US-09-529-206A-49
; Sequence 49, Application US/09529206A
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; FILE REFERENCE: 20264269US1
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US/09/529,206A
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: PCT/US98/19609
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206A-49

Query Match      100.0%; Score 45; DB 19; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.05;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AMLAAQERV 10
   |||||
Db 1 AMLAAQERV 10

RESULT 4
US-09-529-206B-49
; Sequence 49, Application US/09529206B
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; FILE REFERENCE: 20264269US1
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US/09/529,206B
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206B-49

Query Match      100.0%; Score 45; DB 19; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.05;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AMLAAQERV 10
   |||||
Db 1 AMLAAQERV 10

RESULT 5
US-09-529-206D-49
; Sequence 49, Application US/09529206D
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; FILE REFERENCE: 218791
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: PCT/US98/19609
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206D-49

Query Match      100.0%; Score 45; DB 19; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.05;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AMLAAQERV 10
   |||||
Db 1 AMLAAQERV 10

RESULT 6
US-09-529-206-50
; Sequence 50, Application US/09529206
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America as
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; FILE REFERENCE: 20264269PC
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US/09/529,206
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206-50

Query Match      100.0%; Score 45; DB 19; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.05;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AMLAAQERV 10
   |||||
Db 1 AMLAAQERV 10
```

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:25:43 ; Search time 78.7342 Seconds  
(without alignments)  
115.568 Million cell updates/sec

Title: US-09-807-512-25  
Perfect score: 45  
Sequence: 1 AMLAAQERRV 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched:	5728757 seqs, 909918778 residues	5728757
Total number of hits satisfying chosen parameters:		

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 10%
                  Listing first 45

```

```
Database : Pending_Patents_AA_Main:*
1: /cgn2_6/pptodata/1/paa/PCTrus_COMB.pep.*
2: /cgn2_6/pptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/pptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/pptodata/1/paa/US080_COMB.pep.*
5: /cgn2_6/pptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/pptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/pptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/pptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/pptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/pptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/pptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/pptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/pptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/pptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/pptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/pptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/pptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/pptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/pptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/pptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/pptodata/1/paa/US097A_COMB.pep.*
22: /cgn2_6/pptodata/1/paa/US097B_COMB.pep.*
23: /cgn2_6/pptodata/1/paa/US098_COMB.pep.*
24: /cgn2_6/pptodata/1/paa/US099A_COMB.pep.*
25: /cgn2_6/pptodata/1/paa/US099B_COMB.pep.*
26: /cgn2_6/pptodata/1/paa/US100_COMB.pep.*
27: /cgn2_6/pptodata/1/paa/US101_COMB.pep.*
28: /cgn2_6/pptodata/1/paa/US102_COMB.pep.*
29: /cgn2_6/pptodata/1/paa/US103_COMB.pep.*
30: /cgn2_6/pptodata/1/paa/US104_COMB.pep.*
31: /cgn2_6/pptodata/1/paa/US106_COMB.pep.*
32: /cgn2_6/pptodata/1/paa/US107_COMB.pep.*
33: /cgn2_6/pptodata/1/paa/US108_COMB.pep.*
```

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result	Query	Score	Match	Length	DB	ID	Description
1	45	100.0	10	23	US-09-807-512-25		Sequence 25, Appl

2	45	100.0	12	19	US-09-529-206-49	Sequence 49, Appl
3	45	100.0	12	19	US-09-529-206A-49	Sequence 49, Appl
4	45	100.0	12	19	US-09-529-206B-49	Sequence 49, Appl
5	45	100.0	12	19	US-09-529-206D-49	Sequence 49, Appl
6	45	100.0	13	19	US-09-529-206-50	Sequence 50, Appl
7	45	100.0	13	19	US-09-529-206A-50	Sequence 50, Appl
8	45	100.0	13	19	US-09-529-206B-50	Sequence 50, Appl
9	45	100.0	13	19	US-09-529-206D-50	Sequence 50, Appl
10	45	100.0	25	19	US-09-529-206-56	Sequence 56, Appl
11	45	100.0	25	19	US-09-529-206A-56	Sequence 56, Appl
12	45	100.0	25	19	US-09-529-206B-56	Sequence 56, Appl
13	45	100.0	25	19	US-09-529-206D-56	Sequence 56, Appl
14	45	100.0	58	19	US-09-529-206-5	Sequence 5, Appl
15	45	100.0	58	19	US-09-529-206A-5	Sequence 5, Appl
16	45	100.0	58	19	US-09-529-206B-5	Sequence 5, Appl
17	45	100.0	58	19	US-09-529-206D-5	Sequence 5, Appl
18	45	100.0	58	23	US-09-807-512-10	Sequence 10, Appl
19	45	100.0	109	23	US-09-807-512-2	Sequence 2, Appl
20	41	91.1	9	23	US-09-807-512-36	Sequence 26, Appl
21	41	91.1	11	19	US-09-529-206-48	Sequence 48, Appl
22	41	91.1	11	19	US-09-529-206A-48	Sequence 48, Appl
23	41	91.1	11	19	US-09-529-206B-48	Sequence 48, Appl
24	41	91.1	11	19	US-09-529-206D-48	Sequence 48, Appl
25	38	80.0	9	19	US-09-417-507-29765	Sequence 29765, A
26	36	80.0	9	19	US-09-529-206A-123	Sequence 123, App
27	36	80.0	9	19	US-09-529-206B-123	Sequence 123, App
28	36	80.0	9	19	US-09-529-206D-123	Sequence 123, App
29	36	80.0	10	17	US-09-336-091-45	Sequence 45, Appl
30	36	80.0	10	19	US-09-529-206-47	Sequence 47, Appl
31	36	80.0	10	19	US-09-529-206A-47	Sequence 47, Appl
32	36	80.0	10	19	US-09-529-206B-47	Sequence 47, Appl
33	36	80.0	10	19	US-09-529-206D-47	Sequence 47, Appl
34	36	80.0	117	16	US-09-252-691-10675	Sequence 10675, A
35	36	80.0	117	16	US-09-252-691C-10675	Sequence 10675, A
36	36	80.0	117	30	US-10-417-886-10675	Sequence 886, Ap
37	36	80.0	325	27	US-10-155-881-8908	Sequence 8908, Ap
38	35	77.8	182	28	US-09-417-507-44277	Sequence 44277, A
39	35	77.8	259	28	US-10-282-122A-47627	Sequence 47627, A
40	34	75.6	317	30	US-10-425-114-71411	Sequence 71411, A
41	34	75.6	341	38	US-09-428-944-1639	Sequence 1639, Ap
42	34	75.6	369	30	US-10-437-963-137731	Sequence 137731, A
43	34	75.6	450	30	US-10-425-114-68789	Sequence 68789, A
44	34	75.6	551	29	US-10-369-493-20673	Sequence 20673, A
45	34	75.6	551	32	US-60-360-039-20673	Sequence 20673, A

## ALIGNMENTS

```

RESULT 1
US-09-807-512-25
US-09-807-512-25
Sequence 25, Application US/09807512
GENERAL INFORMATION:
;
; APPLICANT: Schrier, Peter I.
; APPLICANT: Rarnoudse, Corlien
; APPLICANT: Heider, Karl-Heinz
; APPLICANT: Klade, Christoph.
; TITLE OF INVENTION: Camel, An Alternative Translation Product of the Tumor
; TITLE OF INVENTION: Antigen-Lage 1
; FILE REFERENCE: 0652.2200000
; CURRENT APPLICATION NUMBER: US/09/807,512
; CURRENT FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: PCT/EP99/07832
; PRIOR FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: EP 98119583.7
; PRIOR FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 10
; TYPE: prt
; ORGANISM: Homo sapiens
US-09-807-512-25

```

**THIS PAGE BLANK (USPTO)**

Search completed: October 7, 2003, 13:58:57  
Job time : 11.8861 secs

---

US-10-156-761-14955

Query Match 68.9%; Score 31; DB 15; Length 202;  
Best Local Similarity 75.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAAQRRV 10  
| | | | | :  
Db 132 LAAQRRQI 139

RESULT 12

US-10-156-761-8217  
; Sequence 8217, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 8217  
; LENGTH: 263  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-8217

Query Match 68.9%; Score 31; DB 15; Length 263;  
Best Local Similarity 66.7%; Pred. No. 1.7e+02;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 MLAAQRRV 10  
: | | | : | | |  
Db 140 LLAARDRV 148

RESULT 13

US-10-156-761-7637  
; Sequence 7637, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 7637  
; LENGTH: 331  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-7637

Query Match 68.9%; Score 31; DB 15; Length 331;  
Best Local Similarity 70.0%; Pred. No. 2.2e+02;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AMLAAQRRV 10  
| | | | | :  
Db 81 AELAAQQRV 90

RESULT 14

US-10-156-761-14852  
; Sequence 14852, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 14852  
; LENGTH: 408  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-14852

Query Match 66.7%; Score 30; DB 15; Length 408;  
Best Local Similarity 77.8%; Pred. No. 4.2e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 MLAAQRRV 10  
| | | | | :  
Db 123 MLAEQARRV 131

RESULT 15

US-09-769-734-10  
; Sequence 10, Application US/09769734  
; Publication No. US20030143666A1  
; GENERAL INFORMATION:  
; APPLICANT: Ecopia BioSciences Inc.  
; TITLE OF INVENTION: Genetic Locus for Everninomicin Biosynthesis  
; FILE REFERENCE: PA 005-US  
; CURRENT APPLICATION NUMBER: US/09/769,734  
; CURRENT FILING DATE: 2001-01-26  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 10  
; LENGTH: 429  
; TYPE: PRT  
; ORGANISM: M. carbonacea  
US-09-769-734-10

Query Match 66.7%; Score 30; DB 12; Length 429;  
Best Local Similarity 70.0%; Pred. No. 4.4e+02;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AMLAAQRRV 10  
| | | | | :  
Db 279 AELAAQEREL 288

```
Db      85 ALLGAEERKV 94

RESULT 7
US-10-000-151B-3
; Sequence 3, Application US/10000151B
; Publication No. US20030013136A1
; GENERAL INFORMATION:
; APPLICANT: Balser, Jeffrey R.
; APPLICANT: George, Alfred L.
; TITLE OF INVENTION: HUMAN KCB1 REGULATION OF HERG POTASSIUM CHANNEL BLOCK
; FILE REFERENCE: Vanderbilt Ref No. US20030013136A1 VU0120; Attorney Docket No. US2003
; CURRENT APPLICATION NUMBER: 2000-10-30
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1159
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-000-151B-3

Query Match      71.1%; Score 32; DB 15; Length 1159;
Best Local Similarity 60.0%; Pred. No. 4.8e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 AMLAAQERRV 10
       1:1 1:1:1:1
Db      85 ALLGAEERKV 94

RESULT 8
US-10-193-692-5
; Sequence 5, Application US/10193692
; Publication No. US20030074682A1
; GENERAL INFORMATION:
; APPLICANT: Exelixis Deutschland GmbH
; TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF A NOVEL TELEOST POTASSIUM
; FILE REFERENCE: AR02-005C
; CURRENT APPLICATION NUMBER: US/10/193,692
; CURRENT FILING DATE: 2002-07-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1159
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-193-692-5

Query Match      71.1%; Score 32; DB 15; Length 1159;
Best Local Similarity 60.0%; Pred. No. 4.8e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 AMLAAQERRV 10
       1:1 1:1:1:1
Db      85 ALLGAEERKV 94

RESULT 9
US-10-185-867-10
; Sequence 10, Application US/10185867
; Publication No. US20030104429A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; TITLE OF INVENTION: NOVEL POTASSIUM CHANNEL MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-055CP
; CURRENT APPLICATION NUMBER: US/10/185,867
; CURRENT FILING DATE: 2002-06-27
; PRIOR APPLICATION NUMBER: US/09/358,383
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: USSN 09/119,855
; PRIOR FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 36

Query Match      71.1%; Score 32; DB 15; Length 1159;
Best Local Similarity 60.0%; Pred. No. 4.8e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 AMLAAQERRV 10
       1:1 1:1:1:1
Db      85 ALLGAEERKV 94

RESULT 10
US-10-174-613-49
; Sequence 49, Application US/10174613
; Publication No. US20030114354A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL POTASSIUM CHANNEL WITH HOMOLOG
; FILE REFERENCE: D0123 NP
; CURRENT APPLICATION NUMBER: US/10/174,613
; CURRENT FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US 60/299,378
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/300,614
; PRIOR FILING DATE: 2001-06-25
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 49
; LENGTH: 1159
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-174-613-49

Query Match      71.1%; Score 32; DB 15; Length 1159;
Best Local Similarity 60.0%; Pred. No. 4.8e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 AMLAAQERRV 10
       1:1 1:1:1:1
Db      85 ALLGAEERKV 94

RESULT 11
US-10-156-761-14955
; Sequence 14955, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14955
; LENGTH: 202
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
```



; FILE REFERENCE: D0123 NP  
; CURRENT APPLICATION NUMBER: US/10/174,613  
; CURRENT FILING DATE: 2002-06-19  
; PRIOR APPLICATION NUMBER: US 60/299,378  
; PRIOR FILING DATE: 2001-06-19  
; PRIOR APPLICATION NUMBER: US 60/300,614  
; PRIOR FILING DATE: 2001-06-25  
; NUMBER OF SEQ ID NOS: 93  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 24  
; LENGTH: 110  
; TYPE: PR1  
; ORGANISM: Homo sapiens  
US-10-174-613-24

Query Match 71.1%; Score 32; DB 15; Length 110;  
Best Local Similarity 60.0%; Pred. No. 45;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 AMLAAQERV 10  
|:| |:|:|  
Db 60 ALLGAERKV 69

## RESULT 3

US-10-156-761-12932  
; Sequence 12932, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:

; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; PRIOR FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 12932  
; LENGTH: 179  
; TYPE: PR1  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-12932

Query Match 71.1%; Score 32; DB 15; Length 179;  
Best Local Similarity 100.0%; Pred. No. 73;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AAQERV 10  
|:| |:|:|  
Db 24 AAQERV 30

## RESULT 4

US-09-735-995-2  
; Sequence 2, Application US/09735995  
; Patent No. US20010034024A1  
; GENERAL INFORMATION:

; APPLICANT: Keating, Mark T.  
; APPLICANT: Splawski, Igor  
; TITLE OF INVENTION: MUTATIONS IN AND GENOMIC STRUCTURE OF HERG - A LONG QT  
; FILE REFERENCE: 2323-136  
; CURRENT APPLICATION NUMBER: US/09/735,995  
; CURRENT FILING DATE: 2000-12-14  
; PRIOR APPLICATION NUMBER: 09/226,012

; PRIOR FILING DATE: 1999-01-06  
; NUMBER OF SEQ ID NOS: 116  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 1159  
; TYPE: PR1  
; ORGANISM: Homo sapiens  
US-09-735-995-2

Query Match 71.1%; Score 32; DB 9; Length 1159;  
Best Local Similarity 60.0%; Pred. No. 4.8e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 AMLAAQERV 10  
|:| |:|:|  
Db 85 ALLGAERKV 94

## RESULT 5

US-09-735-995-4  
; Sequence 4, Application US/09735995  
; Patent No. US20010034024A1  
; GENERAL INFORMATION:

; APPLICANT: Keating, Mark T.  
; APPLICANT: Splawski, Igor  
; TITLE OF INVENTION: MUTATIONS IN AND GENOMIC STRUCTURE OF HERG - A LONG QT  
; FILE REFERENCE: 2323-136  
; CURRENT APPLICATION NUMBER: US/09/735,995  
; CURRENT FILING DATE: 2000-12-14  
; PRIOR APPLICATION NUMBER: 09/226,012  
; PRIOR FILING DATE: 1999-01-06  
; NUMBER OF SEQ ID NOS: 116  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 1159  
; TYPE: PR1  
; ORGANISM: Homo sapiens  
US-09-735-995-4

Query Match 71.1%; Score 32; DB 9; Length 1159;  
Best Local Similarity 60.0%; Pred. No. 4.8e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 AMLAAQERV 10  
|:| |:|:|  
Db 85 ALLGAERKV 94

## RESULT 6

US-09-119-855-10  
; Sequence 10, Application US/09119855  
; Patent No. US20020099197A1  
; GENERAL INFORMATION:

; APPLICANT: Curtiss, Rory A.J.  
; TITLE OF INVENTION: NOVEL POTASSIUM CHANNEL MOLECULES AND USES THEREFOR  
; FILE REFERENCE: mmi-055  
; CURRENT APPLICATION NUMBER: US/09/119,855  
; CURRENT FILING DATE: 1998-07-21  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 10  
; LENGTH: 1159  
; TYPE: PR1  
; ORGANISM: Homo sapiens  
US-09-119-855-10

Query Match 71.1%; Score 32; DB 9; Length 1159;  
Best Local Similarity 60.0%; Pred. No. 4.8e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 AMLAAQERV 10  
|:| |:|:|

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model  
Run on: October 7, 2003, 13:31:20 ; Search time 10.8861 Seconds  
(without alignments)  
145.335 Million cell updates/sec

Title: US-09-807-512-25  
Perfect score: 45  
Sequence: 1 AMLAAQERRV 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues 587654

Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : Published Applications\_AA:\*
- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
  - 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
  - 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
  - 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
  - 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
  - 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
  - 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
  - 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
  - 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
  - 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
  - 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
  - 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	33	73.3	8360	12 US-10-132-134-34	Sequence 34, Appl
2	32	71.1	1110	15 US-10-174-613-24	Sequence 24, Appl
3	32	71.1	179	15 US-10-156-761-12932	Sequence 12932, A
4	32	71.1	1159	9 US-09-735-995-2	Sequence 2, Appl
5	32	71.1	1159	9 US-09-735-995-4	Sequence 4, Appl
6	32	71.1	1159	9 US-09-119-855-10	Sequence 10, Appl
7	32	71.1	1159	15 US-10-000-151B-3	Sequence 3, Appl
8	32	71.1	1159	15 US-10-193-692-5	Sequence 5, Appl
9	32	71.1	1159	15 US-10-185-867-10	Sequence 10, Appl
10	32	71.1	1159	15 US-10-174-613-49	Sequence 49, Appl
11	31	68.9	202	15 US-10-156-761-14955	Sequence 14955, A
12	31	68.9	263	15 US-10-156-761-8217	Sequence 8217, Ap
13	31	68.9	331	15 US-10-156-761-7637	Sequence 7637, Ap
14	30	66.7	408	15 US-10-156-761-14852	Sequence 14852, A
15	30	66.7	429	12 US-09-769-734-10	Sequence 10, Appl

16	30	66.7	652	15 US-10-156-761-14695	Sequence 14695, A
17	30	66.7	758	14 US-10-117-846-16	Sequence 16, Appl
18	30	66.7	819	15 US-10-128-714-3429	Sequence 3429, Ap
19	30	66.7	825	15 US-10-128-714-8429	Sequence 8429, Ap
20	30	66.7	1467	12 US-10-017-161-2044	Sequence 2044, Ap
21	29	64.4	100	11 US-09-764-891-2727	Sequence 2727, Ap
22	29	64.4	153	11 US-09-764-891-4027	Sequence 4027, Ap
23	29	64.4	204	16 US-10-080-170-561	Sequence 561, App
24	29	64.4	282	9 US-09-815-242-5576	Sequence 5576, Ap
25	29	64.4	232	9 US-10-156-761-12440	Sequence 12440, A
26	29	64.4	320	15 US-10-156-761-10213	Sequence 10213, A
27	29	64.4	345	12 US-09-769-734-5	Sequence 5, Appl
28	29	64.4	397	10 US-09-282-879-2	Sequence 2, Appl
29	29	64.4	399	11 US-09-817-513A-2	Sequence 2, Appl
30	29	64.4	431	15 US-10-156-761-8533	Sequence 8533, Ap
31	29	64.4	446	15 US-10-156-761-10790	Sequence 10790, A
32	29	64.4	455	11 US-09-955-999-110	Sequence 110, App
33	29	64.4	521	10 US-09-813-320-4	Sequence 4, Appl
34	29	64.4	530	15 US-09-813-320-2	Sequence 2, Appl
35	29	64.4	636	15 US-10-156-761-12826	Sequence 12826, A
36	29	64.4	942	15 US-10-156-761-10438	Sequence 10438, A
37	29	64.4	980	12 US-10-177-809-10	Sequence 10, Appl
38	29	64.4	1177	15 US-10-193-692-4	Sequence 4, Appl
39	29	64.4	1186	15 US-10-193-692-2	Sequence 2, Appl
40	29	64.4	1195	15 US-10-192-440-10	Sequence 10, Appl
41	29	64.4	1196	15 US-10-255-532-2	Sequence 2, Appl
42	29	64.4	1243	15 US-10-196-935A-4	Sequence 4, Appl
43	29	64.4	1724	10 US-09-964-899-43	Sequence 43, Appl
44	29	64.4	1987	15 US-10-132-382-6	Sequence 6, Appl
45	29	64.4	2013	15 US-10-132-382-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1  
US-10-132-134-34  
; Sequence 34, Application US/10132134  
; Publication No. US20030171562A1  
; GENERAL INFORMATION:  
; APPLICANT: Farnet, Chris  
; APPLICANT: Yang, Xianshu  
; APPLICANT: Staffa, Alfredo  
; APPLICANT: Zazopoulos, Emmanuel  
; TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES  
; FILE REFERENCE: 3012-2US  
; CURRENT APPLICATION NUMBER: US/10/132,134  
; CURRENT FILING DATE: 2002-04-26  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 34  
; LENGTH: 8360  
; TYPE: PRT  
; ORGANISM: Streptomyces amphibiosporus  
US-10-132-134-34

Query Match 73.3%; Score 33; DB 12; Length 8360;  
Best Local Similarity 70.0%; Pred. No. 2.3e+03;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AMLAAQERRV 10  
|:| | | | |  
Db 1748 ALLAAQEQRV 1757

RESULT 2  
US-10-174-613-24  
; Sequence 24, Application US/10174613  
; Publication No. US20030114354A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL POTASSIUM CHANNEL WITH HOMOLOG  
; TITLE OF INVENTION: Ether-a-go-go FAMILY, HEAG2

**THIS PAGE BLANK (USPTO)**

```

: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
: ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
:
: NUMBER OF SEQUENCES: 7310
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: GENOME THERAPEUTICS CORPORATION
: STREET: 100 Beaver Street
: CITY: Waltham
: STATE: Massachusetts
: COUNTRY: USA
: ZIP: 02354
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: CD-ROM ISO9660
: COMPUTER: PC
: OPERATING SYSTEM: <Unknown>
: SOFTWARE: ASCII
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/107,532A
: FILING DATE: 30-Jun-1998
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 60/085,598
: FILING DATE: 14 May 1998
: APPLICATION NUMBER: 60/051571
: FILING DATE: July 2, 1997
:
: ATTORNEY/AGENT INFORMATION:
: NAME: Ariniello, Pamela Deneke
: REGISTRATION NUMBER: 40,489
: REFERENCE/DOCKET NUMBER: GTC-012
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (781)893-5007
: TELEFAX: (781)893-8277
:
: INFORMATION FOR SEQ ID NO: 6921:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 207 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: HYPOTHETICAL: YES
: ORIGINAL SOURCE:
: ORGANISM: Enterococcus faecium
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (B) LOCATION 1...207
:
: SEQUENCE DESCRIPTION: SEQ ID NO: 6921:
US-09-107-532A-6921

```

```

Query Match      64.4%; Score 29; DB 4; Length 207;
Best Local Similarity 75.0%; Pred. NO. 1.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 AMLAAQER 8
      ||| |||
Db     121 AMIAEQER 128

```

```

Search completed: October 7, 2003, 13:34:07
Job time : 5.44304 secs

```

; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 5205  
; LENGTH: 427  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-5205

Query Match 66.7%; Score 30; DB 4; Length 427;  
Best Local Similarity 66.7%; Pred. No. 1.6e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AMLAAQERR 9  
| : ||||| :  
Db 333 ATIAAQEKR 341

RESULT 12  
US-09-252-991A-27953  
; Sequence 27953, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 27953  
; LENGTH: 451  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-27953

Query Match 66.7%; Score 30; DB 4; Length 451;  
Best Local Similarity 66.7%; Pred. No. 1.7e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 MLAAQERRV 10  
: : |||||  
Db 259 LVIAQERRV 267

RESULT 13  
US-09-252-991A-31942  
; Sequence 31942, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 31942  
; LENGTH: 662  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-31942

Query Match 66.7%; Score 30; DB 4; Length 662;  
Best Local Similarity 75.0%; Pred. No. 2.6e+02;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
Qy 2 MLAAQERR 9  
: ||||| :  
Db 222 VLAAEERR 229

RESULT 14  
US-09-107-532A-6626  
; Sequence 6626, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD/ROM ISO9660  
; COMPUTER: PC  
; OPERATING SYSTEM: <Unknown>  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,532A  
; FILING DATE: 30-Jun-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/085,598  
; FILING DATE: 14 May 1998  
; APPLICATION NUMBER: 60/051571  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Arinello, Pamela Deneke  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTC-012  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (781)893-5007  
; TELEFAX: (781)893-8277  
; INFORMATION FOR SEQ ID NO: 6626:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 89 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; ORIGINAL SOURCE:  
; ORGANISM: Enterococcus faecium  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (B) LOCATION 1...89  
; SEQUENCE DESCRIPTION: SEQ ID NO: 6626:  
US-09-107-532A-6626  
Query Match 64.4%; Score 29; DB 4; Length 89;  
Best Local Similarity 75.0%; Pred. No. 48;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 AMLAAQER 8  
| : ||| :  
Db 3 AMIAQER 10

RESULT 15  
US-09-107-532A-6921  
; Sequence 6921, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-358-383C-10

Query Match      71.1%; Score 32; DB 4; Length 1159;
Best Local Similarity 60.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 AMLAAQERRV 10
Db      85 ALLGAERKV 94

RESULT 7
US-09-252-991A-20779
; Sequence 20779, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20779
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20779

Query Match      68.9%; Score 31; DB 4; Length 122;
Best Local Similarity 70.0%; Pred. No. 27;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 AMLAAQERRV 10
Db      47 AMLTSVERRV 56

RESULT 8
US-09-252-991A-27383
; Sequence 27383, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27383
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27383

Query Match      66.7%; Score 30; DB 4; Length 303;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 AMLAAQERR 9
Db      213 AMLAVQOOR 221
```

```
RESULT 9
US-09-252-991A-28531
; Sequence 28531, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28531
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28531

Query Match      66.7%; Score 30; DB 4; Length 330;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 AMLAAQERRV 10
Db      78 ALLAQHRRRL 87

RESULT 10
US-09-252-991A-24774
; Sequence 24774, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24774
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24774

Query Match      66.7%; Score 30; DB 4; Length 380;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 AMLAAQERR 9
Db      48 AVVAAQERQ 56

RESULT 11
US-09-328-352-5205
; Sequence 5205, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACT;
; FILE REFERENCE: GTC99-03PA
```

```
Query Match      75.6%; Score 34; DB 1; Length 103;
Best Local Similarity 80.0%; Pred. No. 5.5;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AMLAAQERRV 10
Db 28 AAXAAQERRV 37

RESULT 2
US-08-956-242-13
; Sequence 13, Application US/08956242C
; Patent No. 5986081
; GENERAL INFORMATION:
; APPLICANT: Ganetzký, Barry S.
; TITLE OF INVENTION: Polynucleotides Encoding Herg-3
; FILE REFERENCE: 960296.94550
; CURRENT APPLICATION NUMBER: US/08/956,242C
; CURRENT FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 1159
; TYPE: PRT
; ORGANISM: Homo sapien
US-08-956-242-13

Query Match      71.1%; Score 32; DB 2; Length 1159;
Best Local Similarity 60.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 AMLAAQERRV 10
Db 85 ALLGAERKV 94

RESULT 3
US-09-351-215-13
; Sequence 13, Application US/09351215
; Patent No. 6087488
; GENERAL INFORMATION:
; APPLICANT: Ganetzký, Barry S.
; APPLICANT: Titus, Steven A.
; TITLE OF INVENTION: Polynucleotides Encoding Herg-3
; FILE REFERENCE: 960296.94550
; CURRENT APPLICATION NUMBER: US/09/351,215
; CURRENT FILING DATE: 1999-07-12
; EARLIER APPLICATION NUMBER: 08/956,242
; EARLIER FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 1159
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-351-215-13

Query Match      71.1%; Score 32; DB 3; Length 1159;
Best Local Similarity 60.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 AMLAAQERRV 10
Db 85 ALLGAERKV 94

RESULT 4
US-09-226-012-2
; Sequence 2, Application US/09226012
; Patent No. 6207383
; GENERAL INFORMATION:
; APPLICANT: Keating, Mark T.
; APPLICANT: Splawski, Igor
; TITLE OF INVENTION: MUTATIONS IN AND GENOMIC STRUCTURE OF HERG - A LONG QT
; FILE REFERENCE: 2323-136
; CURRENT APPLICATION NUMBER: US/09/226,012
; CURRENT FILING DATE: 1999-01-06
; EARLIER APPLICATION NUMBER: 09/122,847
; EARLIER FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1159
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-226-012-2

Query Match      71.1%; Score 32; DB 3; Length 1159;
Best Local Similarity 60.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 AMLAAQERRV 10
Db 85 ALLGAERKV 94

RESULT 5
US-09-226-012-4
; Sequence 4, Application US/09226012
; Patent No. 6207383
; GENERAL INFORMATION:
; APPLICANT: Keating, Mark T.
; APPLICANT: Splawski, Igor
; TITLE OF INVENTION: MUTATIONS IN AND GENOMIC STRUCTURE OF HERG - A LONG QT
; FILE REFERENCE: 2323-136
; CURRENT APPLICATION NUMBER: US/09/226,012
; CURRENT FILING DATE: 1999-01-06
; EARLIER APPLICATION NUMBER: 09/122,847
; EARLIER FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1159
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-226-012-4

Query Match      71.1%; Score 32; DB 3; Length 1159;
Best Local Similarity 60.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 AMLAAQERRV 10
Db 85 ALLGAERKV 94

RESULT 6
US-09-358-383C-10
; Sequence 10, Application US/09358383C
; Patent No. 6518398
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; TITLE OF INVENTION: NOVEL POTASSIUM CHANNEL MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-055CP
; CURRENT APPLICATION NUMBER: US/09/358,383C
; CURRENT FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: USSN 09/119,855
; PRIOR FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 1159
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:21:07 ; search time 5.44304 Seconds  
(without alignments)  
77.734 Million cell updates/sec

Title: US-09-807-512-25  
Perfect score: 45  
Sequence: 1 AMLAQERRV 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/PCT05\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	75.6	103	1 US-08-552-142A-13	Sequence 13, Appl
2	32	71.1	1159	2 US-08-956-242-13	Sequence 13, Appl
3	32	71.1	1159	3 US-09-351-215-13	Sequence 13, Appl
4	32	71.1	1159	3 US-09-226-012-2	Sequence 2, Appl
5	32	71.1	1159	3 US-09-226-012-4	Sequence 4, Appl
6	32	71.1	1159	4 US-09-358-383C-10	Sequence 10, Appl
7	31	68.9	122	4 US-09-252-991A-20779	Sequence 20779, A
8	30	66.7	303	4 US-09-252-991A-27383	Sequence 27383, A
9	30	66.7	330	4 US-09-252-991A-28531	Sequence 28531, A
10	30	66.7	380	4 US-09-252-991A-24774	Sequence 24774, A
11	30	66.7	427	4 US-09-328-352-5205	Sequence 5205, Ap
12	30	66.7	451	4 US-09-252-991A-27953	Sequence 27953, A
13	30	66.7	662	4 US-09-252-991A-31942	Sequence 31942, A
14	29	64.4	89	4 US-09-107-532A-6626	Sequence 6626, Ap
15	29	64.4	207	4 US-09-107-532A-6921	Sequence 6921, Ap
16	29	64.4	210	4 US-09-107-532A-6445	Sequence 6445, Ap
17	29	64.4	263	4 US-08-708-573F-10	Sequence 10, Appl
18	29	64.4	268	4 US-08-708-573F-14	Sequence 14, Appl
19	29	64.4	272	4 US-09-252-991A-24539	Sequence 24539, A
20	29	64.4	302	4 US-09-252-991A-25971	Sequence 25971, A
21	29	64.4	311	4 US-09-252-991A-22295	Sequence 22295, A
22	29	64.4	321	4 US-09-252-991A-18807	Sequence 18807, A
23	29	64.4	333	1 US-08-118-270-12	Sequence 12, Appl
24	29	64.4	333	5 PCT-US93-08528-12	Sequence 12, Appl
25	29	64.4	397	2 US-08-774-104A-2	Sequence 2, Appl
26	29	64.4	399	2 US-08-839-581A-2	Sequence 2, Appl
27	29	64.4	399	3 US-09-023-591A-2	Sequence 2, Appl

## ALIGNMENTS

## RESULT 1

US-08-552-142A-13  
; Sequence 13, Application US/08552142A  
; Patent No. 5695995  
; GENERAL INFORMATION:  
; APPLICANT: Weintraub, Harold M.  
; APPLICANT: Lee, Jacqueline E.  
; APPLICANT: Tapscott, Stephen J.  
; APPLICANT: Hollenberg, Stanley M.  
; TITLE OF INVENTION: Neurogenic Differentiation (NeuroD) Genes  
; TITLE OF INVENTION: and Proteins  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Christensen O'Connor Johnson KindnessPLLC  
; STREET: 1420 Fifth Avenue, Suite 2800  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101-2347  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/552.142A  
; FILING DATE: 02-NOV-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/239,238  
; FILING DATE: 06-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/US95/05741  
; FILING DATE: 08-MAY-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Broderick, Thomas F.  
; REGISTRATION NUMBER: 31,332  
; REFERENCE/DOCKET NUMBER: FHCR-1-8933  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 206-682-8100  
; TELEFAX: 206-225-0709  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 103 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens

US-08-552-142A-13

Sequence 5, Appl  
Sequence 23070, A  
Sequence 25723, A  
Sequence 4462, Ap  
Sequence 2, Appl  
Sequence 6, Appl  
Sequence 7, Appl  
Sequence 3, Appl  
Sequence 6, Appl  
Sequence 7, Appl  
Sequence 29, Appl  
Sequence 307, App  
Sequence 309, App  
Sequence 1390, Ap  
Sequence 306, App  
Sequence 10, Appl  
Sequence 11, Appl

28 29 64.4 429 4 US-09-351-150A-5  
29 29 64.4 472 4 US-09-252-991A-23070  
30 29 64.4 575 4 US-09-252-991A-25723  
31 29 64.4 593 4 US-09-328-352-4462  
32 29 64.4 1088 3 US-09-082-059-6  
33 28 62.2 34 1 US-08-129-089-6  
34 28 62.2 34 1 US-08-129-089-7  
35 28 62.2 34 1 US-08-454-444-2  
36 28 62.2 34 1 US-08-454-444-3  
37 28 62.2 34 5 PCT-US93-01135-6  
38 28 62.2 34 5 PCT-US93-01135-7  
39 28 62.2 58 4 US-09-358-383C-29  
40 28 62.2 136 4 US-09-732-210-307  
41 28 62.2 137 4 US-09-732-210-309  
42 28 62.2 137 4 US-09-732-210-1390  
43 28 62.2 149 4 US-09-732-210-306  
44 28 62.2 150 1 US-08-362-453-10  
45 28 62.2 150 1 US-08-362-453-11



**THIS PAGE BLANK (USPTO)**

DT 01-JUL-1997 (TREMBLrel. 04, Created)  
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Lysostaphin immunity factor.  
 GN LIF.  
 OS Staphylococcus simulans.  
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 OX NCBI\_TaxID=1286;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97260121; PubMed=9106216;  
 RA Thumm G., Gotz F.;  
 RT "Studies on polysostaphin processing and characterization of the  
 RT lysostaphin immunity factor (Lif) of Staphylococcus simulans biovar  
 RT staphyloolyticus".  
 RL Mol. Microbiol. 23:1251-1265(1997).  
 DR EMBL; U66883; AAB53784.1; -.  
 DR InterPro; IPR003447; Meth\_resist.  
 DR Pfam; PF02388; FemAB; 1.  
 SQ SEQUENCE 413 AA; 49121 MW; B183532C6D96F551 CRC64;

Query Match 71.1%; Score 32; DB 2; Length 413;  
 Best Local Similarity 75.0%; Pred. No. 2e+02; Indels 0; Gaps 0;  
 Matches 6; Conservative 2; Mismatches 0;  
 Qy 3 LAAQRRV 10  
 Db 289 LAAQERKI 296  
 |||||:

RESULT 14  
 Q9BUC6 PRELIMINARY; PRT; 456 AA.  
 ID Q9BUC6 PRELIMINARY; PRT; 456 AA.  
 AC Q9BUC6;  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE Similar to cell cycle progression 2 protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Tissue-Uterus;  
 RA Strausberg R.;  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC002732; AA02732.1; -.  
 SQ SEQUENCE 456 AA; 51368 MW; 19D4B1AC5A47BA75 CRC64;

Query Match 71.1%; Score 32; DB 4; Length 456;  
 Best Local Similarity 87.5%; Pred. No. 2.2e+02; Indels 0; Gaps 0;  
 Matches 7; Conservative 0; Mismatches 1;  
 Qy 2 MAAQERR 9  
 Db 89 MAAQSR 96  
 |||||:

RESULT 15  
 Q8NDM4 PRELIMINARY; PRT; 631 AA.  
 ID Q8NDM4 PRELIMINARY; PRT; 631 AA.  
 AC Q8NDM4;  
 DT 01-OCT-2002 (TREMBLrel. 22, Created)  
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)  
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
 DE Hypothetical protein.  
 GN DKFZP727A151.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC Tissue-Breast Cancer;  
 RA Poustka A., Klein M., Mewes H.W., Gassenhuber J., Wiemann S.;  
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL833840; CAD38700.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 631 AA; 70770 MW; 1199508B6CBEB03 CRC64;  
 Query Match 71.1%; Score 32; DB 4; Length 631;  
 Best Local Similarity 87.5%; Pred. No. 3e+02; Indels 0; Gaps 0;  
 Matches 7; Conservative 0; Mismatches 1;  
 Qy 2 MAAQERR 9  
 Db 264 MAAQSR 271  
 |||||:

Search completed: October 7, 2003, 13:31:04  
 Job time : 19.7089 secs

```

RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Uterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AE005851; AAK23747.1; -.
DR TIGR; CC1771; -.
DR InterPro; IPR000298; CytC_oxdse_III.
DR Pfam; PF00510; COX3; 1.
DR ProDom; PD000382; CytC_oxdse_III; 1.
DR PROSITE; PS0253; COX3; 1.
KW Complete proteome.
SQ SEQUENCE 208 AA; 22965 MW; F371176CE6AE0649 CRC64;

Query Match 71.1%; Score 32; DB 16; Length 208;
Best Local Similarity 70.0%; Pred. No. 99;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AMLAAQERRV 10
Db 89 AMIAQAQREV 98
||:|||||
|:|||||

RESULT 10
Q9L126 PRELIMINARY; PRT; 208 AA.
AC Q9L126;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Putative RNA polymerase sigma factor.
GN SCO2639 OR SC8E4A.09C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939113; CAB71814.1; -.
KW Complete proteome.
SQ SEQUENCE 208 AA; 22978 MW; A535200CE7DC603D CRC64;

Query Match 71.1%; Score 32; DB 16; Length 208;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AAQERRV 10
Db 53 AAQERRV 59
||:|||||
|:|||||

RESULT 11
O14710 PRELIMINARY; PRT; 360 AA.
AC O14710;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)

```

```

DE Cell cycle progression 2 protein.
GN CPR2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98043401; PubMed=9383053;
RA Edwards M.C., Liegeois N., Horecka J., DePinho R.A., Sprague G.F. Jr.,
RA Tvers M., Elledge S.J.;
RT "Human CPR (cell cycle progression restoration) genes impart a Far-
RT phenotype on yeast cells.";
RL Genetics 147:1063-1076(1997).
DR EMBL; AF011792; AAB69312.1; -.
SQ SEQUENCE 360 AA; 40254 MW; 9AC0F8E46DA11199 CRC64;

Query Match 71.1%; Score 32; DB 4; Length 360;
Best Local Similarity 87.5%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MLAAQERR 9
Db 37 MLAAQSRR 44
||:|||||
|:|||||

RESULT 12
O19059 PRELIMINARY; PRT; 404 AA.
AC O19059;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE 11 beta-hydroxysteroid dehydrogenase type 2 (Fragment).
GN 11<BETA>-HSD2.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96386030; PubMed=8793860;
RA Campbell L.E., Yu M., Yang K.;
RT "Ovine 11 beta-hydroxysteroid dehydrogenase type 2 gene predicts a
RT protein distinct from that deduced by the cloned kidney cDNA at the C-
RT terminus.";
RL Mol. Cell. Endocrinol. 119:113-118(1996).
CC -!- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
DR EMBL; S83516; AAB50810.1; -.
DR EMBL; S83532; AAB50810.1; JOINED.
DR HSPSP; P14061; LFDS.
DR InterPro; IPR002198; ADH_short.
DR Pfam; PF00106; adh_short; 1.
DR PROSITE; PS00061; ADH_SHORT; 1.
KW Oxidoreductase.
FT NON_TER
FT NON_TER
SQ SEQUENCE 404 AA; 43978 MW; F34A6600D7D1E6D5 CRC64;

Query Match 71.1%; Score 32; DB 6; Length 404;
Best Local Similarity 77.8%; Pred. No. 1.9e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AMLAAQERR 9
Db 328 ALLAAQPRR 336
||:|||||
|:|||||

RESULT 13
O05989 PRELIMINARY; PRT; 413 AA.
ID O05989
AC O05989;

```

DB	84	VLAQAQERR 91	
RESULT 6			
Q9A618			
ID	Q9A618	PRELIMINARY;	PRT; 306 AA.
AC	Q9A618;		
DT	01-JUN-2001 (TrEMBLrel. 17, Created)		
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)		
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)		
DE	Glycosyl transferase family protein.		
GN	CC2277.		
OS	Caulobacter crescentus.		
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;		
OC	Caulobacteraceae; Caulobacter.		
OX	NCBI_TaxID=155892;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-ATCC 19089 / CB15;		
RX	MEDLINE-21173698; PubMed=11259647;		
RA	Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,		
RA	Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,		
RA	Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,		
RA	DeBoy R.T., Dodson R.J., Durkin A.S., Ginn M.L., Haft D.H.,		
RA	Kolony J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,		
RA	Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,		
RA	Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;		
RT	"Complete genome sequence of Caulobacter crescentus.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).		
DR	EMBL; AE005898; AK24248.1; -		
DR	TIGR; CC2277; -		
DR	InterPro; IPR001173; Glyco_trans_2.		
DR	Pfam; PF00535; Glycos_transf_2; 1.		
KW	Transferase; Complete proteome.		
SQ	SEQUENCE 306 AA; 33352 MW; 77F25076EB8A62D4 CRC64;		
Query Match	73.3%;	Score 33;	DB 16; Length 306;
Best Local Similarity	87.5%;	Pred. No. 90;	
Matches	7; Conservative	1; Mismatches	0; Indels 0; Gaps 0;
QY	1 AMLAAQER 8		
DB	105 ALLAAQER 112		
RESULT 7			
Q8EI18			
ID	Q8EI18	PRELIMINARY;	PRT; 350 AA.
AC	Q8EI18;		
DT	01-MAR-2003 (TrEMBLrel. 23, Created)		
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)		
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)		
DE	Nicotinate-nucleotide--dimethylbenzimidazole		
DE	phosphoribosyltransferase.		
GN	COBT OR SO1035.		
OS	Shewanella oneidensis.		
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;		
OC	Alteromonadaceae; Shewanella.		
OX	NCBI_TaxID=70863;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-MR-1;		
RX	MEDLINE-22297686; PubMed=12368813;		
RA	Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,		
RA	Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,		
RA	Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,		
RA	DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,		
RA	Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,		
RA	Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,		
RA	Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,		
RA	Feldblyum T.V., Smith H.O., Venter J.C., Nealeson K.H., Fraser C.M.;		
RT	"Genome sequence of the dissimilatory metal ion-reducing bacterium		
RT	Shewanella oneidensis.";		
RL	Nat. Biotechnol. 20:1118-1123(2002).		
DR	EMBL; AE015548; AAN54108.1; -		
DR	TIGR; SO1035; -		
KW	Complete proteome.		
SQ	SEQUENCE 350 AA; 36784 MW; 4C152362503A532E CRC64;		
Query Match	73.3%;	Score 33;	DB 16; Length 350;
Best Local Similarity	70.0%;	Pred. No. 1e+02;	
Matches	7; Conservative	2; Mismatches	1; Indels 0; Gaps 0;
QY	1 AMLAAQERR 10		
DB	249 AMLAAAEKRM 258		
RESULT 8			
Q8Y139			
ID	Q8Y139	PRELIMINARY;	PRT; 169 AA.
AC	Q8Y139;		
DT	01-MAR-2002 (TrEMBLrel. 20, Created)		
DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)		
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)		
DE	Hypothetical protein RSC0854.		
GN	RSC0854 OR RS04967.		
OS	Ralstonia solanacearum (Pseudomonas solanacearum).		
OC	Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;		
OC	Ralstoniaceae; Ralstonia.		
OX	NCBI_TaxID=305;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-GMI1000;		
RX	MEDLINE-21681879; PubMed=11823852;		
RA	Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,		
RA	Ariat M., Billault A., Brottier P., Camus J.C., Cattolico L.,		
RA	Chandler M., Choisme N., Claudel-Renard C., Cunac S., Demange N.,		
RA	Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,		
RA	Sigulier P., Thebault P., Whalen M., Wincker P., Levy M.,		
RA	Weissenbach J., Boucher C.A.;		
RT	"Genome sequence of the plant pathogen Ralstonia solanacearum.";		
RL	Nature 415:497-502(2002).		
DR	EMBL; AL646061; CAD14556.1; -		
KW	Hypothetical protein; Complete proteome.		
SQ	SEQUENCE 169 AA; 17682 MW; 7EC99E253BF84F92 CRC64;		
Query Match	71.1%;	Score 32;	DB 16; Length 169;
Best Local Similarity	70.0%;	Pred. No. 81;	
Matches	7; Conservative	2; Mismatches	1; Indels 0; Gaps 0;
QY	1 AMLAAQERR 10		
DB	114 ALLAAVDRRV 123		
RESULT 9			
Q9A7F2			
ID	Q9A7F2	PRELIMINARY;	PRT; 208 AA.
AC	Q9A7F2;		
DT	01-JUN-2001 (TrEMBLrel. 17, Created)		
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)		
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)		
DE	Ubiquinol oxidase subunit III.		
GN	CC1771.		
OS	Caulobacter crescentus.		
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;		
OC	Caulobacteraceae; Caulobacter.		
OX	NCBI_TaxID=155892;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-ATCC 19089 / CB15;		
RX	MEDLINE-21173698; PubMed=11259647;		
RA	Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,		
RA	Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,		
RA	Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,		
RA	DeBoy R.T., Dodson R.J., Durkin A.S., Ginn M.L., Haft D.H.,		
RA	Kolony J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,		
RA	Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,		
RA	Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;		
RT	"Complete genome sequence of Caulobacter crescentus.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).		
DR	EMBL; AE005898; AK24248.1; -		
DR	TIGR; CC2277; -		
DR	InterPro; IPR001173; Glyco_trans_2.		
DR	Pfam; PF00535; Glycos_transf_2; 1.		
KW	Transferase; Complete proteome.		
SQ	SEQUENCE 306 AA; 33352 MW; 77F25076EB8A62D4 CRC64;		
Query Match	73.3%;	Score 33;	DB 16; Length 306;
Best Local Similarity	87.5%;	Pred. No. 90;	
Matches	7; Conservative	1; Mismatches	0; Indels 0; Gaps 0;
QY	1 AMLAAQER 8		
DB	105 ALLAAQER 112		

QY 1 AMLAAQERRV 10  
|:|:|:|:|:|  
Db 16 AMLAAQERRV 25

## RESULT 2

ID O95987 PRELIMINARY; PRT; 109 AA.  
AC O95987;  
DT 01-MAY-1999 (T-EMBLrel. 10, Created)  
DT 01-MAY-1999 (T-EMBLrel. 10, Last sequence update)  
DT 01-MAY-1999 (T-EMBLrel. 10, Last annotation update)  
DE CTL-recognized antigen ON melanoma (CAMEL).  
GN CAMEL.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Melanoma;  
RA Aarnoudse C.A., Doel vanden P.B., Heemskerk B., Schrier P.I.;  
RT "IL-2 induced melanoma-specific CTL recognize CAMEL, an unexpected  
translation product of LAGE-1";  
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ012835; CAA10197.1; -;  
DR EMBL; AJ012833; CAA10193.1; -;  
DR EMBL; AJ012834; CAA10195.1; -;  
SQ SEQUENCE 109 AA; 11689 MW; 188F85BC04C1F5F0 CRC64;

Query Match 100.0%; Score 45; DB 4; Length 109;  
Best Local Similarity 100.0%; Pred. No. 0.099;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AMLAAQERRV 10  
|:|:|:|:|:|  
Db 16 AMLAAQERRV 25

## RESULT 3

ID Q8W424 PRELIMINARY; PRT; 336 AA.  
AC Q8W424;  
DT 01-MAR-2002 (T-EMBLrel. 20, Created)  
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)  
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)  
DE Hypothetical 37.7 kDa protein (Fragment).  
GN F14.  
OS Gossypium hirsutum (Upland cotton).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.  
OX NCBI\_TaxID=3635;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. CRI 12; TISSUE=Fiber;  
RA Zhao G.R., Liu J.Y.;  
RT "Identification and expression of mRNA transcripts in growing cotton  
fiber cells by fluorescent differential display";  
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ311665; CAC84115.1; -;  
DR InterPro; IPR005027; Glyco\_transf\_43.  
DR Pfam; PF03360; Glyco\_transf\_43; 1.  
FT NON\_TER  
KW Hypothetical protein.  
SQ SEQUENCE 336 AA; 37675 MW; 78AD75C27BBD07AC CRC64;

Query Match 75.6%; Score 34; DB 10; Length 336;  
Best Local Similarity 70.0%; Pred. No. 61;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AMLAAQERRV 10  
|:|:|:|:|:|

Db 205 ALLSANERRV 214

## RESULT 4

ID Q8X0T0 PRELIMINARY; PRT; 655 AA.  
AC Q8X0T0;  
DT 01-MAR-2002 (T-EMBLrel. 20, Created)  
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)  
DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)  
DE Hypothetical 71.6 kDa protein.  
GN 18F11.080.  
OS Neurospora crassa.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariales; Sordariaceae; Neurospora.  
OX NCBI\_TaxID=5141;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,  
RA Nyakatura G., Mewes H.W., Mannhaupt G.;  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA German Neurospora genome project;  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL670011; CAD21413.1; -;  
KW Hypothetical protein.  
SQ SEQUENCE 655 AA; 71619 MW; 159ADE04A519425F CRC64;

Query Match 75.6%; Score 34; DB 3; Length 655;  
Best Local Similarity 87.5%; Pred. No. 1.2e+02;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAAQERRV 10  
|:|:|:|:|:|  
Db 265 IAAQERRV 272

## RESULT 5

ID Q8H8L8 PRELIMINARY; PRT; 129 AA.  
AC Q8H8L8;  
DT 01-MAR-2003 (T-EMBLrel. 23, Created)  
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)  
DE Hypothetical protein.  
GN OSJNBA0070N04.25.  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzeae; Oryza.  
OX NCBI\_TaxID=39947;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Nipponbare;  
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,  
RA Overton II L.L., Tsitrin T., Kim M.M., Bera J.J., Jin S.S.,  
RA Fardosh D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,  
RA Yanag S.S., Riedmuller S.B., Utterback T.T., Feldblym T.V.,  
RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,  
RA White O., Salzberg S.L., Fraser C.M.;  
RT "Oryza sativa chromosome 3 BAC OSJNBA0070N04 genomic sequence.";  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC091494; AAN65022.1; -;  
KW Hypothetical protein.  
SQ SEQUENCE 129 AA; 13913 MW; 073ED4F410DE5D51 CRC64;

Query Match 73.3%; Score 33; DB 10; Length 129;  
Best Local Similarity 87.5%; Pred. No. 38;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MLAAQERRV 9  
|:|:|:|:|:|

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:18:33 ; Search time 16.7089 Seconds  
(without alignments)  
154.441 Million cell updates/sec

Title: US-09-807-512-25  
Perfect score: 45  
Sequence: 1 AMLAAQERRV 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues  
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_23.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_virus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	45	100.0	58	O95146	O95146 homo sapien
2	45	100.0	109	O95987	O95987 homo sapien
3	34	75.6	336	O8W424	O8W424 gossypium h
4	34	75.6	655	O8X0T0	O8X0T0 neurospora
5	33	73.3	129	O8H8L8	O8H8L8 oryza sativ
6	33	73.3	306	O8A618	O8A618 caulobacter
7	33	73.3	350	O8E118	O8E118 shewanella
8	32	71.1	169	O8Y139	O8Y139 ralsstonia s
9	32	71.1	208	O9A7F2	O9A7F2 caulobacter
10	32	71.1	268	O9L126	O9L126 streptomyce
11	32	71.1	360	O14710	O14710 homo sapien
12	32	71.1	404	O19059	O19059 ovis aries
13	32	71.1	413	O05989	O05989 staphylococ
14	32	71.1	456	O9BUC6	O9BUC6 homo sapien
15	32	71.1	631	O8NDM4	O8NDM4 homo sapien
16	32	71.1	631	O96920	O96920 homo sapien

17	32	71.1	801	9	Q9FZT0	Q9fzt0 pseudomonas
18	32	71.1	829	10	Q8LIY9	Q8liy9 oryza sativ
19	32	71.1	1058	4	Q8IZI2	Q8izi2 homo sapien
20	32	71.1	6889	16	Q8XS40	Q8xs40 ralsstonia s
21	31	68.9	31	11	Q9QXB6	Q9qxb6 mus musculus
22	31	68.9	205	16	Q9RD77	Q9rd77 streptomyce
23	31	68.9	215	2	Q9WX74	Q9wx74 acetobacteri
24	31	68.9	242	16	O53794	O53794 mycobacteri
25	31	68.9	277	2	Q9JPD1	Q9jpd1 rhodocyclus
26	31	68.9	293	16	Q91065	Q91065 pseudomonas
27	31	68.9	339	11	Q9IX26	Q9ix26 mus musculus
28	31	68.9	339	11	Q9QXD7	Q9qxd7 mus musculus
29	31	68.9	368	2	Q9ALR3	Q9alr3 pseudomonas
30	31	68.9	383	16	Q9A2S7	Q9a2s7 caulobacter
31	31	68.9	401	17	Q9HS28	Q9hs28 halobacteri
32	31	68.9	495	2	P72346	P72346 synechococc
33	31	68.9	497	16	Q93JB9	Q93jb9 streptomyce
34	31	68.9	500	16	Q9A3K8	Q9a3k8 caulobacter
35	31	68.9	650	5	Q9W472	Q9w472 drosophila
36	31	68.9	1047	5	Q22985	Q22985 caenorhabdi
37	31	68.9	4801	2	Q9AIT3	Q9ait3 xanthomonas
38	30	66.7	96	16	Q92P91	Q92p91 rhizobium m
39	30	66.7	163	2	Q8RJW3	Q8rjw3 xanthomonas
40	30	66.7	165	17	O58344	O58344 pyrococcus
41	30	66.7	168	17	O9UYR8	O9uyr8 pyrococcus
42	30	66.7	179	2	O8VS40	O8vs40 klebsiella
43	30	66.7	194	16	Q8G6B9	Q8g6b9 bifidobacte
44	30	66.7	233	16	Q9S2R6	Q9s2r6 streptomyce
45	30	66.7	235	16	Q9RTZ7	Q9rtz7 deinococcus

ALIGNMENTS

RESULT 1.

O95146	PRELIMINARY;	PRT;	58 AA.
ID	O95146;		
AC	O95146;		
DT	01-MAY-1999 (TREMBLrel. 10, Created)		
DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)		
DT	01-MAR-2001 (TREMBLrel. 16, Last annotation update)		
DE	LAGE-2ALT protein isoform.		
GN	LAGE-2.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Wang R.-F., Johnston S.L., Zeng G., Topalian S.L.,		
RA	Schwarzentruber D.J., Rosenberg S.A.;		
RT	"A Breast and Melanoma-Shared Tumor Antigen: T Cell Responses to		
RT	Antigenic Peptides Translated from Different Open Reading Frames.";		
RL	J. Immunol. 161:3596-3606(1998).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RA	Lethe B.G.;		
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=99454989; PubMed=10523621;		
RA	De Smet C., Turquin C., Lethe B., Martelange V., Boon T.;		
RT	"DNA methylation is the primary silencing mechanism for a set of germ		
RT	line- and tumor-specific genes with a CpG-rich promoter.";		
RL	Mol. Cell. Biol. 19:7327-7335(1999).		
DR	EMBL; AF038567; AAD05203.1; -		
DR	EMBL; AJ275977; CAB76944.1; -		
SQ	SEQUENCE 58 AA; 6188 MW; ED12057564BC7EF2 CRC64;		

Query Match 100.0%; Score 45; DB 4; Length 58;  
Best Local Similarity 100.0%; Pred. No. 0.053;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; AE000998; AAB89738.1; -  
DR PIR; A69438; A69438.  
DR TIGR; AF1506; -  
DR InterPro; IPR005676; Asp\_ADH.f.  
DR InterPro; IPR000319; Asp\_Semiald\_dh.  
DR InterPro; IPR000534; Semiald\_dh.  
DR Pfam; PF01118; Semialdehyde\_dh; 1.  
DR Pfam; PF02774; Semialdehyde\_dhc; 1.  
DR TIGRFAMS; TIGR00978; asd\_EA; 1.  
DR PROSITE; PS01103; ASD; 1.  
KW Oxidoreductase; NADP; Methionine biosynthesis; Threonine biosynthesis;  
KW Amino-acid biosynthesis; Complete proteome.  
FT ACT\_SITE 148 148 BY SIMILARITY.  
SQ SEQUENCE 343 AA; 37696 MW; A6C9F5D045CBFC21 CRC64;

Query Match 68.9%; Score 31; DB 1; Length 343;  
Best Local Similarity 87.5%; Pred. No. 35;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 LAAQERV 10  
   | | | | |  
Db 33 LAASERV 40

Search completed: October 7, 2003, 13:26:27  
Job time : 3.91139 secs

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC DR EMBL; U39722; AAC71631.1; -;  
CC DR PIR; F64244; F64244.  
CC DR TIGR; MG403; -;  
CC DR InterPro; IPR005864; ATPaseB.  
CC DR InterPro; IPR002146; ATPsynth\_B/B'sub.  
CC DR Pfam; PF00430; ATP-synt\_B; 1  
CC DR TIGRFAMs; TIGR01144; ATP-synt\_b; 1.  
CC DR PROSITE; PS00013; PROKAR\_LIPOPROTEIN; 1.  
CC DR Hydrogen ion transport; Transmembrane; CF(0); Membrane; Lipoprotein;  
CC KW Signal; Complete proteome.  
CC FT SIGNAL 1 27  
CC FT CHAIN 28 208  
CC FT LIPID 28 28  
CC FT TRANSMEM 49 69  
CC FT SEQUENCE 208 AA; 24602 MW; 0AB6526A367F84E6 CRC64;  
CC  
CC Query Match 68.9%; Score 31; DB 1; Length 208;  
CC Best Local Similarity 100.0%; Pred. No. 21;  
CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CC  
CC QY 1 AMLAAQE 7  
CC DB 172 AMLAAQE 178  
CC  
CC RESULT 15  
CC DHAS\_ARCFU  
CC ID DHAS\_ARCFU STANDARD; PRT; 343 AA.  
CC AC O28766;  
CC DT 30-MAY-2000 (Rel. 39, Created)  
CC DT 30-MAY-2000 (Rel. 39, Last sequence update)  
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)  
CC DE Aspartate-semialdehyde dehydrogenase (EC 1.2.1.11) (ASA  
CC dehydrogenase) (ASADH).  
CC GN ASD OR AF1506.  
CC OS Archaeoglobus fulgidus.  
CC OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;  
CC OC Archaeoglobaceae; Archaeoglobus.  
CC OX NCBI\_TaxID=2234;  
CC [1]  
CC RP SEQUENCE FROM N.A.  
CC RX STRAIN=VC-16 / DSM 4304 / ATCC 49558;  
CC RX MEDLINE=98049343; PubMed=9389475;  
CC RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,  
CC RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,  
CC RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,  
CC RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,  
CC RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,  
CC RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,  
CC RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,  
CC RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,  
CC RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,  
CC RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,  
CC RA Venter J.C.;  
CC RT "The complete genome sequence of the hyperthermophilic,  
CC sulphate-reducing archaeon Archaeoglobus fulgidus.";  
CC RL Nature 390:364-370(1997).  
CC CC -!- FUNCTION: THIS ENZYME CATALYZES THE SECOND STEP IN THE COMMON  
CC METABOLIC PATHWAY TO SYNTHESIZE THR AND MET FROM ASPARTIC ACID.  
CC CC -!- CATALYTIC ACTIVITY: L-aspartate-4-semialdehyde + phosphate +  
CC NADP(+) = L-4-aspartyl phosphate + NADPH.  
CC CC -!- PATHWAY: METHIONINE BIOSYNTHESIS, THREONINE BIOSYNTHESIS.  
CC CC -!- SIMILARITY: Belongs to the aspartate-semialdehyde dehydrogenase  
CC family.  
CC -----

CC Pfam; PF00027; cNMP\_binding; 1.  
CC DR PIR; F64244; F64244.  
CC DR TIGR; MG403; -;  
CC DR InterPro; IPR005864; ATPaseB.  
CC DR InterPro; IPR002146; ATPsynth\_B/B'sub.  
CC DR Pfam; PF00430; ATP-synt\_B; 1  
CC DR TIGRFAMs; TIGR01144; ATP-synt\_b; 1.  
CC DR PROSITE; PS00013; PROKAR\_LIPOPROTEIN; 1.  
CC DR Hydrogen ion transport; Transmembrane; CF(0); Membrane; Lipoprotein;  
CC KW Signal; Complete proteome.  
CC FT SIGNAL 1 27  
CC FT CHAIN 28 208  
CC FT LIPID 28 28  
CC FT TRANSMEM 49 69  
CC FT SEQUENCE 208 AA; 24602 MW; 0AB6526A367F84E6 CRC64;  
CC  
CC Query Match 71.1%; Score 32; DB 1; Length 1163;  
CC Best Local Similarity 60.0%; Pred. No. 70;  
CC Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
CC  
CC QY 1 AMLAAQERV 10  
CC DB 85 ALLGAERKV 94  
CC  
CC RESULT 14  
CC ATPF\_MYCGE  
CC ID ATPF\_MYCGE STANDARD; PRT; 208 AA.  
CC AC P47643;  
CC DT 01-FEB-1996 (Rel. 33, Created)  
CC DT 01-FEB-1996 (Rel. 33, Last sequence update)  
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)  
CC DE ATP synthase B chain precursor (EC 3.6.3.14).  
CC GN ATPF OR MG403.  
CC OS Mycoplasma genitalium.  
CC OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
CC OX NCBI\_TaxID=2097;  
CC [1]  
CC RP SEQUENCE FROM N.A.  
CC RX STRAIN=ATCC 33530 / G-37;  
CC RX MEDLINE=96026346; PubMed=7569993;  
CC RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,  
CC RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,  
CC RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,  
CC RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,  
CC RA Tomb J.-F., Dougherty B.A., Bitt K.F., Hu P.-C., Lucier T.S.,  
CC RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;  
CC RT "The minimal gene complement of Mycoplasma genitalium.";  
CC RL Science 270:397-403(1995).  
CC CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +  
CC H(+) (Out).  
CC CC -!- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC  
CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE  
CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)  
CC HAS THREE MAIN SUBUNITS: A, B AND C (By similarity).  
CC CC -!- SIMILARITY: BELONGS TO THE ATPASE B CHAIN FAMILY.  
CC -----



KW Potassium channel; Potassium transport; Transmembrane;  
KW Phosphorylation; Glycoprotein; Multigene family; Polymorphism;  
KW Alternative splicing.

FT DOMAIN 1 405  
FT TRANSMEM 406 426  
FT TRANSMEM 453 473  
FT DOMAIN 474 497  
FT TRANSMEM 498 518  
FT TRANSMEM 523 543  
FT DOMAIN 544 549  
FT TRANSMEM 550 570  
FT DOMAIN 614 634  
FT TRANSMEM 641 661  
FT DOMAIN 662 1162  
FT DOMAIN 17 88  
FT DOMAIN 92 144  
FT DOMAIN 299 302  
FT NP\_BIND 744 861  
FT CARBOHYD 600 600  
FT VARSPLIC 1 59  
FT VARSPLIC 1 378

FT FT 186 186  
FT VARIANT 455 455  
FT VARIANT 752 752  
FT VARIANT 1006 1006  
FT CONFLICT 111 111  
FT CONFLICT 126 126  
FT CONFLICT 198 198  
FT CONFLICT 214 214  
FT CONFLICT 537 537  
FT CONFLICT 908 908  
FT CONFLICT 929 929  
FT SEQUENCE 1162 AA; 126885 MW; A9455F7F10B61E46 CRC64;

Query Match 71.1%; Score 32; DB 1; Length 1162;  
Best Local Similarity 60.0%; Pred. No. 70;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 AMLAAQERV 10  
Db 85 ALLGAEKRV 94  
i:1|:11|:1

RESULT 13  
KCH2\_RAT STANDARD; PRT; 1163 AA.  
AC O08962; O08720;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Potassium voltage-gated channel subfamily H member 2 (Ether-a-go-go  
DE related gene potassium channel 1) (ERG1) (r-ERG) (RERG) (Ether-a-go-  
DE go related protein 1) (Eag related protein 1).  
GN KCNH2 OR ERG.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain cortex;

RX MEDLINE=98329322; PubMed=9664620;  
RA Bauer C.K., Engeland B., Wulfsen I., Ludwig J., Pongs O.,  
RA Schwarz J.R.;  
RT "RERG is a molecular correlate of the inward-rectifying K current in  
RT clonal net pituitary cells.";  
RL Recept. Channels 6:19-29(1998).  
RN [2]  
RP SEQUENCE OF 409-568 FROM N.A.  
RX MEDLINE=97164986; PubMed=9012748;  
RA Wymore R.S., Gintant G.A., Wymore R.T., Dixon J.E., McKinnon D.,  
RA Cohen I.S.;  
RT "Tissue and species distribution of mRNA for the IKr-like K+ channel,  
RT erg.";  
RL Circ. Res. 80:261-268(1997).  
RN [3]  
RX TISSUE SPECIFICITY.  
RX MEDLINE=20183472; PubMed=10718922;  
RA Wulfsen I., Hauber H.P., Schiemann D., Bauer C.K., Schwarz J.R.;  
RT "Expression of mRNA for voltage-dependent and inward-rectifying K  
RT channels in GH3/B6 cells and rat pituitary.";  
RL J. Neuroendocrinol. 12:263-272(2000).  
RN [4]  
RP INTERACTION WITH KCNH6 AND KCNH7, AND MUTAGENESIS OF GLY-630.  
RX MEDLINE=21079731; PubMed=11212207;  
RA Wimmers S., Wulfsen I., Bauer C.K., Schwarz J.R.;  
RT "Erg1, erg2 and erg3 K channel subunits are able to form  
RT heteromultimers.";  
RL Pflugers Arch. 441:450-455(2001).  
CC -1- FUNCTION: Pore-forming (alpha) subunit of voltage-gated inwardly  
CC rectifying potassium channel. Channel properties are modulated by  
CC cAMP and subunit assembly. Mediates the rapidly activating  
CC component of the delayed rectifying potassium current in heart  
CC (IKr) (By similarity).  
CC -1- SUBUNIT: The potassium channel is probably composed of a homo- or  
CC heterotetrameric complex of pore-forming alpha subunits that can  
CC associate with modulating beta subunits. Heteromultimer with  
CC KCNH6/ERG2 and KCNH7/ERG3. Heteromultimer with  
CC KCNE1 and KCNE2 (By similarity).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -1- TISSUE SPECIFICITY: Highly expressed in brain and testis, slightly  
CC less so in heart, adrenal, retina and thymus. Detected at lower  
CC levels in lung, soleus, tibialis, and at very low levels in cornea  
CC and lens. A shorter transcript is detected in skeletal muscle.  
CC Found in pituitary.  
CC -1- DOMAIN: The segment S4 is probably the voltage-sensor and is  
CC characterized by a series of positively charged amino acids at  
CC every third position.  
CC -1- PTM: Phosphorylated on serine and threonine residues (By  
CC similarity).  
CC -1- SIMILARITY: Belongs to the potassium channel family. H (Eag)  
CC subfamily.  
CC -1- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.  
CC -1- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.  
CC -1- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; Z96106; CAB09536.1; -;  
DR EMBL; U75210; AAC53160.1; -;  
DR InterPro; IPR000595; cNMP\_binding.  
DR InterPro; IPR003967; Erg\_Channel.  
DR InterPro; IPR005821; Ion\_Trans.  
DR InterPro; IPR001622; K+channel\_pore.  
DR InterPro; IPR005820; M+channel\_nlg.  
DR InterPro; IPR001610; PAC.  
DR InterPro; IPR000700; PAS-associat.  
DR InterPro; IPR000014; PAS\_domain.

```

FT TRANSMEM 406 426 SEGMENT S1 (POTENTIAL).
FT TRANSMEM 433 473 SEGMENT S2 (POTENTIAL).
FT DOMAIN 474 497 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 498 518 SEGMENT S3 (POTENTIAL).
FT TRANSMEM 523 543 SEGMENT S4 (POTENTIAL).
FT DOMAIN 544 549 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 550 570 SEGMENT S5 (POTENTIAL).
FT DOMAIN 614 634 SEGMENT H5 (PORE-FORMING) (POTENTIAL).
FT TRANSMEM 641 661 SEGMENT S6 (POTENTIAL).
FT DOMAIN 662 1161 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 17 88 PAC.
FT DOMAIN 92 144 POLY-PRO.
FT DOMAIN 299 302 CNMP.
FT NP_BIND 744 861 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 600 600 Missing (in isoform 2).
FT VARSPLIC 69 85 /FTIG-VSP-000971.
FT CONFLICT 411 411 V -> A (IN REF. 2).
FT CONFLICT 445 446 PE -> TD (IN REF. 2).
FT CONFLICT 553 553 L -> F (IN REF. 2).
FT CONFLICT 561 561 L -> C (IN REF. 2).
SQ SEQUENCE 1161 AA; 126961 MW; 79B532B2FFBD9AEB CRC64;

Query Match 71.1%; Score 32; DB 1; Length 1161;
Best Local Similarity 60.0%; Pred No. 70;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 AMLAQERV 10
|:|:|:|:|
Db 85 ALLGAERKV 94

RESULT 12
KCH2_MOUSE STANDARD; PRT; 1162 AA.
AC O35219; O35220; O35221; O35989;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Potassium voltage-gated channel subfamily H member 2 (Ether-a-go-go
DE related gene potassium channel 1) (ERGI) (MERG) (Merg1) (Ether-a-go-
DE go related protein 1) (Eag related protein 1).
GN KCHN2 OR ERG OR MERG1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3), AND VARIANTS ARG-186;
RP THR-455; TVR-752 AND ASN-1006.
RC STRAIN=129/Sv, and BALB/c; TISSUE=Heart;
RX MEDLINE=98012815; PubMed=9351462;
RA London B., Trudeau M.C., Newton K.P., Beyer A.K., Copeland N.G.,
RA Gilbert D.J., Jenkins N.A., Satter C.A., Robertson G.A.;
RA "Two isoforms of the mouse ether-a-go-go-related gene coassemble to
RT form channels with properties similar to the rapidly activating
RT component of the cardiac delayed rectifier K+ current.";
RT Circ. Res. 81:870-878(1997).
RL [2]
RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).
RP TISSUE=atrial tumor;
RC MEDLINE=98012799; PubMed=9351446;
RX Lees-Miller J.P., Kondo C., Wang L., Duff H.J.;
RA "Electrophysiological characterization of an alternatively processed
RT ERG K+ channel in mouse and human hearts.";
RL Circ. Res. 81:719-726(1997).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANT ARG-186.
RC TISSUE=Colon smooth muscle;
RA Shoen F., Malykhina A., Akbarali H.I.;
RA "Smooth muscle KCNH2 erg potassium channel.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -|- FUNCTION: Pore-forming (alpha) subunit of voltage-gated inwardly

```

```

CC rectifying potassium channel. Channel properties are modulated by
CC CAMP and subunit assembly. Mediates the rapidly activating
CC component of the delayed rectifying potassium current in heart
CC (IKr) (By similarity).
CC -|- SUBUNIT: The potassium channel is probably composed of a homo- or
CC heterotetrameric complex of pore-forming alpha subunits that can
CC associate with modulating beta subunits. Heteromultimer with
CC KCNH6/ERG2, KCNH7/ERG3, KCNE1 and KCNE2 (By similarity).
CC -|- SUBCELLULAR LOCATION: Integral membrane protein.
CC -|- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Comment=Experimental confirmation may be lacking for some
CC isoforms;
CC Name=1; Synonyms=1A, A;
CC IsoId=O35219-1; Sequence=Displayed;
CC Name=2; Synonyms=1A';
CC IsoId=O35219-2; Sequence=VSP_000969;
CC Name=3; Synonyms=1B, B;
CC IsoId=O35219-3; Sequence=VSP_000970;
CC -|- TISSUE SPECIFICITY: Isoform 1 is expressed in heart, brain and
CC testis and at low levels in lung. Isoform 3 is expressed
CC predominantly in heart. The expression of isoform 2 is low in all
CC tissues tested.
CC -|- DOMAIN: The segment S4 is probably the voltage-sensor and is
CC characterized by a series of positively charged amino acids at
CC every third position.
CC -|- PTM: Phosphorylated on serine and threonine residues (By
CC similarity).
CC -|- SIMILARITY: Belongs to the potassium channel family. H (Eag)
CC subfamily.
CC -|- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.
CC -|- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
CC -|- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
CC -|- CAUTION: Ref.3 sequence was originally reported as deriving from
CC rabbit.
CC -|- CAUTION: Ref.3 sequence differs from that shown due to a
CC frameshift in position 1057.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF012868; AAC53418.1; -
CC EMBL; AF012869; AAC53419.1; -
CC EMBL; AF012871; AAC53420.1; -
CC EMBL; AF012870; AAC53420.1; JOINED.
CC EMBL; AF012871; AAC53421.1; -
CC EMBL; AF012871; AAC53421.1; -
CC EMBL; AF012870; AAC53422.1; -
CC EMBL; AF012870; AAC53422.1; JOINED.
CC EMBL; AF034762; AAB87571.1; -
CC EMBL; AF439342; AAL35327.2; ALT_FRAME.
CC MGD; MGI:1341722; Kcnh2.
CC InterPro; IPR000595; CNMP-binding.
CC InterPro; IPR003967; Erg_channel.
CC InterPro; IPR005821; Ion_trans.
CC InterPro; IPR001622; K+channel_pore.
CC InterPro; IPR005820; M+channel_nlg.
CC InterPro; IPR001610; PAC.
CC InterPro; IPR000700; PAS-assoc.C.
CC InterPro; IPR000014; PAS_domain.
CC Pfam; PF00027; CNMP_binding; 1.
CC Pfam; PF00520; Ion_trans; 1.
CC Pfam; PF00785; PAC; 1.
CC SMART; SM00100; CNMP; 1.
CC SMART; SM00086; PAC; 1.
CC PROSITE; PS50042; CNMP_BINDING_3; 1.
CC PROSITE; PS50112; PAS; 1.
CC PROSITE; PS50113; PAC; 1.
CC Transport; Ion transport; Ionic channel; Voltage-gated channel;

```



DE related protein 1) (Bag related protein 1) (eag homolog).  
GN KCNH2 OR HERG OR HERG1 OR ERG OR ERG1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=Heart;  
RX MEDLINE=94211879; PubMed=8159766;  
RA Warnke J.W., Ganetzky B.;  
RT "A family of potassium channel genes related to eag in Drosophila and mammals.";  
RL Proc. Natl. Acad. Sci. U.S.A. 91:3438-3442(1994).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANT LQ72 CVS-534.  
RX MEDLINE=98260867; PubMed=9600240;  
RA Itoh T., Tanaka T., Nagai R., Kamiya T., Sawayama T., Nakayama T., Tomoike H., Sakurada H., Yazaki Y., Nakamura Y.;  
RT "Genomic organization and mutational analysis of HERG, a gene responsible for familial long QT syndrome.";  
RL Hum. Genet. 102:435-439(1998).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RA Downie D., Chapman C.G., Punia P., Rice S., Bahmani F., Murdock P., Pearson N., Randall A.D., Meadows H.J.;  
RT "Potent inhibition of HERG K<sup>+</sup> channels by the neuroprotective agent Slipatrigine.";  
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A. (ISOFORM 3).  
RC TISSUE=Heart;  
RX MEDLINE=21269186; PubMed=11374908;  
RA Soejima H., Kawamoto S., Akai J., Miyoshi O., Arai Y., Morohka T., Matsuo S., Nikiawa N., Kimura A., Okubo K., Mukai T.;  
RT "Isolation of novel heart-specific genes using BodyMap database.";  
RL Genomics 74:115-120(2001).  
RN [5]  
RP SEQUENCE FROM N.A. (ISOFORM 2).  
RA Crociani O., Guasti L., Balzi M., Becchetti A., Wanke E., Olivetto M., Wyomere R.S., Arcangeli A.;  
RT "Cell cycle-dependent expression of HERG1 and HERG1b isoforms in tumor cells.";  
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE OF 1-376 FROM N.A. (ISOFORM 2).  
RC TISSUE=Heart;  
RX MEDLINE=98012815; PubMed=9351462;  
RA London B., Trudeau M.C., Newton K.P., Beyer A.K., Copeland N.G., Gilbert D.J., Jenkins N.A., Satler C.A., Robertson G.A.;  
RT "Two isoforms of the mouse ether-a-go-go-related gene coassemble to form channels with properties similar to the rapidly activating component of the cardiac delayed rectifier K<sup>+</sup> current.";  
RL Circ. Res. 81:870-878(1997).  
RN [7]  
RP SEQUENCE OF 1-376 FROM N.A. (ISOFORM 2).  
RC TISSUE=Heart atrium;  
RX MEDLINE=98012795; PubMed=9351446;  
RA Lees-Miller J.P., Kondo C., Wang L., Duff H.J.;  
RT "Electrophysiological characterization of an alternatively processed ERG K<sup>+</sup> channel in mouse and human hearts.";  
RL Circ. Res. 81:719-726(1997).  
RN [8]  
RP SEQUENCE OF 27-1159 FROM N.A. (ISOFORM 1).  
RA Yang P., Paulussen A., Verhasselt P., Crabbe R., Luyten W., Armstrong M.;  
RT "Analysis of the human HERG gene: intron localisation and identification of a novel inherited mutation associated with long QT.";  
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
RN [9]  
RP SEQUENCE OF 795-1159 FROM N.A. (ISOFORM 3).  
RC TISSUE=Heart ventricle;

RX MEDLINE=98438490; PubMed=9765245;  
RA Kupersmidt S., Snyders D.J., Raes A., Roden D.M.;  
RT "A K<sup>+</sup> channel splice variant common in human heart lacks a C-terminal domain required for expression of rapidly activating delayed rectifier current.";  
RL J. Biol. Chem. 273:27231-27235(1998).  
RN [10]  
RP SEQUENCE OF 59-1159 FROM N.A. (ISOFORM 4), AND SEQUENCE OF 711-1159 FROM N.A. (ISOFORM 1/2).  
RC TISSUE=Brain;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Collins E.A., Grouse L.H., Derge J.G., Klausner R.D., Fellings F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Ditschenko L., Marusina K., Farmer A.A., Rubin G.W., Hong L., Brownstein M.J., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E., Stapleton M., Udwin T.B., Toshiyuki S., Carninci P., Prange C., Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Bouffard G.G., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E., Schnerker A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [11]  
RP MUTAGENESIS OF ASN-598; ASN-629 AND SER-631, AND N-GLYCOSYLATION. MEDLINE=22057172; PubMed=12063277;  
RX Gong Q., Anderson C.L., January C.T., Zhou Z.;  
RA "Role of glycosylation in cell surface expression and stability of HERG potassium channels.";  
RT Am. J. Physiol. 283:H77-H84(2002).  
RN [12]  
RP MUTAGENESIS OF SER-283; SER-890; THR-895 AND SER-1137, AND PHOSPHORYLATION. MEDLINE=20299343; PubMed=10837251;  
RX Cui J., Melman Y., Palma E., Fishman G.I., McDonald T.V.;  
RA "Cyclic AMP regulates the HERG K(+) channel by dual pathways.";  
RL Curr. Biol. 10:671-674(2000).  
RN [13]  
RP INTERACTION WITH KCNE1. MEDLINE=97373956; PubMed=9230439;  
RX McDonald T.V., Yu Z., Ming Z., Palma E., Meyers M.B., Wang K.-W., Goldstein S.A., Fishman G.I.;  
RT "A minK-HERG complex regulates the cardiac potassium current I(Kr)." Nature 388:289-292(1997).  
RN [14]  
RP INTERACTION WITH KCNE2. MEDLINE=99235979; PubMed=10219239;  
RX Abbott G.W., Sesti F., Splawski I., Buck M.E., Lehmann M.H., Timothy K.W., Keating M.T., Goldstein S.A.;  
RT "MiRP1 forms IKr potassium channels with HERG and is associated with cardiac arrhythmia." Cell 97:175-187(1999).  
RN [15]  
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 1-135, AND MUTAGENESIS OF PHE-29 AND TYR-43. MEDLINE=99059500; PubMed=9845367;  
RX Morais Cabral J.H., Lee A., Cohen S.L., Chait B.T., Li M., Mackinnon R.;  
RT "Crystal structure and functional analysis of the HERG potassium channel N terminus: a eukaryotic PAS domain." Cell 95:649-655(1998).  
RN [16]  
RP VARIANTS LQ72 ASP-470; VAL-561 AND SER-628. MEDLINE=95196272; PubMed=7889573;  
RX Curran M.E., Splawski I., Timothy K.W., Vincent G.M., Green E.D.,

```

FT CONFLICT 511 511 K -> R (IN REF. 3).
SQ SEQUENCE 582 AA; 61587 MW; 47A71B57B45AFDD9 CRC64;

Query Match 71.1%; Score 32; DB 1; Length 582;
Best Local Similarity 77.8%; Pred. No. 36;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AMLAAQERR 9
Db 425 ALSAAQERR 433
      1: |||||
      425 ALSAAQERR 433

RESULT 9
KCH2_CANFA STANDARD; PRT; 1158 AA.
AC QTS23; 002719; O18820;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Potassium voltage-gated channel subfamily H member 2 (Ether-a-go-go
DE related gene potassium channel 1) (ERG1) (c-erg) (DERG) (Ether-a-go-
DE go related protein 1) (Eag related protein 1).
GN KCNH2 OR ERG OR CERG.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=21310885; PubMed=11417212;
RA Zehelein J., Zhang W., Koenen M., Graf M., Heinemann S.H., Katus H.A.;
RT "Molecular cloning and expression of CERG, the ether a go-go-related
RT gene from canine myocardium."
RL Pflugers Arch. 442:188-191(2001).
RN [2]
RP SEQUENCE OF 407-566 FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=97164986; PubMed=9012748;
RA Wymore R.S., Gintant G.A., Wymore R.T., Dixon J.E., McKinnon D.,
RA Cohen I.S.;
RT "Tissue and species distribution of mRNA for the IKr-like K+ channel,
RT erg."
RN [3]
RP Circ. Res. 80:261-268(1997).
RX SEQUENCE OF 616-714 FROM N.A.
RC TISSUE=Heart atrium;
RX MEDLINE=99221626; PubMed=10205145;
RA Yue L., Melnyk P., Gaspo R., Wang Z., Nattel S.;
RT "Molecular mechanisms underlying ionic remodeling in a dog model of
RT atrial fibrillation."
RL Circ. Res. 84:776-784(1999).
CC -1- FUNCTION: Pore-forming (alpha) subunit of voltage-gated inwardly
CC rectifying potassium channel. Channel properties are modulated by
CC cAMP and subunit assembly. Mediates the rapidly activating
CC component of the delayed rectifying potassium current in heart
CC (IKr) (By similarity).
CC -1- SUBUNIT: The potassium channel is probably composed of a homo- or
CC heterotrimeric complex of pore-forming alpha subunits that can
CC associate with modulating beta subunits. Heteromultimer with
CC KCNH6/ERG2, KCNH7/ERG3, KCNE1 and KCNE2 (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Highly expressed in left and right atria of
CC the heart, in cortex and hippocampus; detected at intermediate
CC levels in left and right ventricle, Purkinje fibers, cerebellum,
CC thalamus and basal ganglia; detected at low levels in liver,
CC spleen and kidney.
CC -1- DOMAIN: The segment S4 is probably the voltage-sensor and is
CC characterized by a series of positively charged amino acids at
CC every third position.
CC -1- PTM: Phosphorylated on serine and threonine residues (By
CC similarity).
CC -1- SIMILARITY: Belongs to the potassium channel family. H (Eag)

```

```

subfamily.
-1- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.
-1- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
-1- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AJ243344; CAB64868.1; -
EMBL; U75213; AAC48722.1; -
EMBL; AF017429; AAB70524.1; -
InterPro: IPR000595; cNMP_binding.
InterPro: IPR003967; Erg_channel.
InterPro: IPR005821; Ion_trans.
InterPro: IPR001622; K+channel_pore.
InterPro: IPR005820; M+channel_nlg.
InterPro: IPR001610; PAC.
InterPro: IPR000700; PAS-assoc_C.
Pfam: PF00027; cNMP_binding; 1.
Pfam: PF00520; Ion_trans; 1.
Pfam: PF00785; PAC; 1.
SMART; SM00100; cNMP; 1.
SMART; SM00091; PAS; 1.
PROSITE; PS50042; cNMP_BINDING_3; 1.
PROSITE; PS50112; PAC; 1.
PROSITE; PS50113; PAC; 1.
KW Transport; Ion transport; Ionic channel; Voltage-gated channel;
KW Potassium channel; Potassium; Glycoprotein; Multigene family.
FT DOMAIN 1 403 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 403 423 SEGMENT S1 (POTENTIAL).
FT TRANSMEM 450 470 SEGMENT S2 (POTENTIAL).
FT DOMAIN 471 494 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 495 515 SEGMENT S3 (POTENTIAL).
FT TRANSMEM 520 540 SEGMENT S4 (POTENTIAL).
FT DOMAIN 541 545 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 547 567 SEGMENT S5 (POTENTIAL).
FT DOMAIN 611 631 SEGMENT H5 (PORE-FORMING) (POTENTIAL).
FT TRANSMEM 638 658 SEGMENT S6 (POTENTIAL).
FT DOMAIN 659 1158 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 17 88 PAS.
FT DOMAIN 92 144 PAC.
FT DOMAIN 296 299 POLY-PRO.
FT NP_BIND 741 858 CNMP.
FT CARBOHYD 597 597 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 442 442 P -> T (IN REF. 2).
SQ SEQUENCE 1158 AA; 126644 MW; 53C849032B4AA3D0 CRC64;

Query Match 71.1%; Score 32; DB 1; Length 1158;
Best Local Similarity 60.0%; Pred. No. 70;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 AMLAAQERRV 10
      1: |||||
Db 85 ALIGAERKV 94

RESULT 10
KCH2_HUMAN STANDARD; PRT; 1159 AA.
AC Q12809; O75418; O75680; Q9BT72; Q9H3P0;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Potassium voltage-gated channel subfamily H member 2 (Ether-a-go-go
DE related gene potassium channel 1) (H-ERG) (Erg1) (Ether-a-go-go)

```



QY 1 AMLAAQERR 9  
 Db 48 ALSAAQERR 56

## RESULT 5

FTSQ\_RHIME STANDARD; PRT; 309 AA.  
 AC 030993;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cell division protein ftSQ homolog.  
 GN FTSQ OR R02170 OR SMC01872.  
 OS Rhizobium meliloti (Sinorhizobium meliloti).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.  
 OX NCBI\_TaxID=382;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1021;  
 RX MEDLINE=98012980; PubMed=9352931;  
 RA Ma X., Sun Q., Wang R., Singh G., Jonietz E.L., Margolin W.;  
 RT "Interactions between heterologous FtsA and FtsZ proteins at the FtsZ  
 ring.";  
 RL J. Bacteriol. 179:6788-6797(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1021;  
 RX MEDLINE=21396507; PubMed=11481430;  
 RA Capela D., Barloy-Hubier F., Gouzy J., Bothe G., Ampe F., Batut J.,  
 RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,  
 RA Gohrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,  
 RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ransperger U.,  
 RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.;  
 RT "Analysis of the chromosome sequence of the legume symbiont  
 Sinorhizobium meliloti strain 1021.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).  
 CC -1- FUNCTION: THIS PROTEIN MAY BE INVOLVED IN SEPTUM FORMATION (BY  
 CC SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. INNER MEMBRANE  
 CC (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE FTSQ FAMILY.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF024660; AAC45822.1; ALT\_INIT.  
 DR EMBL; AL591789; CAC46749.1;  
 DR InterPro; IPR005548; FtsQ.  
 DR Pfam; PF03799; FtsQ; 1.  
 CC Cell division; Septation; Transmembrane; Inner membrane;  
 CC Complete proteome.  
 KW DOMAIN 1 52 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 53 74 POTENTIAL.  
 FT DOMAIN 75 309 PERIPLASMIC (POTENTIAL).  
 SQ SEQUENCE 309 AA; 34205 MW; 7DCF48DC3AF923D5 CRC64;  
 Query Match 71.1%; Score 32; DB 1; Length 309;  
 Best Local Similarity 66.7%; Pred. No. 19;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 MLAAQERR 10  
 Db 301 MLAAQERR 309

## RESULT 6

DH12\_SHEEP STANDARD; PRT; 427 AA.  
 AC P50168;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Corticosteroid 11-beta-dehydrogenase, isozyme 2 (EC 1.1.1.146) (11-  
 DE DH2) (11-beta-hydroxysteroid dehydrogenase type 2) (11-beta-HSD2)  
 DE (NAD-dependent 11-beta-hydroxysteroid dehydrogenase).  
 GN HSD11B2 OR HSD11K.  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 OX NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=95014413; PubMed=7929304;  
 RA Agarwal A.K., Mune T., Monder C., White P.C.;  
 RT "NAD(+)-dependent isozyme of 11 beta-hydroxysteroid dehydrogenase.  
 RT Cloning and characterization of cDNA from sheep kidney.";  
 RL J. Biol. Chem. 269:25959-25962(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=96065316; PubMed=7588402;  
 RA Agarwal A.K., Mune T., Monder C., White P.C.;  
 RT "Cloning of cDNA encoding an NAD(+)-dependent isozyme of 11 beta-  
 RT hydroxysteroid dehydrogenase in sheep kidney.";  
 RL Endocr. Res. 21:389-397(1995).  
 CC -1- FUNCTION: Catalyzes the conversion of cortisol to the inactive  
 CC metabolite cortisone. Modulates intracellular glucocorticoid  
 CC levels, thus protecting the nonselective mineralocorticoid  
 CC receptor from occupation by glucocorticoids.  
 CC -1- CATALYTIC ACTIVITY: An 11-beta-hydroxysteroid + NAD(+) -> an  
 CC 11-oxosteroid + NADH.  
 CC -1- SUBUNIT: Interacts with ligand-free cytoplasmic NR3C2 (By  
 CC similarity).  
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.  
 CC -1- TISSUE SPECIFICITY: Highly expressed in the kidney and adrenal and  
 CC at lower levels in the colon.  
 CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases  
 CC (SDR) family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U14128; AAA93156.1; ALT\_SEQ.  
 DR PIR; A55353; A55353.  
 DR HSSP; P14061; IFDS.  
 DR InterPro; IPR002198; ADH\_short.  
 DR Pfam; PF00106; adh\_short; 1.  
 DR PROSITE; PS00061; ADH\_SHORT; 1.  
 KW Oxidoreductase; NAD; Mitosome.  
 FT NP\_BIND 82 111 NAD (BY SIMILARITY).  
 FT ACT\_SITE 232 232 BY SIMILARITY.  
 SQ SEQUENCE 427 AA; 46668 MW; 4641749BE408E2B1 CRC64;

Query Match 71.1%; Score 32; DB 1; Length 427;  
 Best Local Similarity 77.8%; Pred. No. 26;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AMLAAQERR 9  
 Db 328 ALLAAQPRR 336

```

Query Match          71.11%; Score 32; DB 1; Length 190;
Best Local Similarity 60.0%; Pred. No. 12;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 AMLAAQERRV 10
      |: || |||:
DB      141 ALFAADERRI 150

RESULT 4
CEAB_ECOLI      STANDARD;      PRT;      205 AA.
ID      P09882;
AC      01-MAR-1989 (Rel. 10, Created)
DT      01-MAR-1989 (Rel. 10, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Colicin E8 (EC 3.1.-.) (Fragment).
COL.
GN      Escherichia coli.
OS      Plasmid ColB8.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC      Enterobacteriaceae; Escherichia.
ON      NCBI_taxid=562;
RX      [1]
SEQUENCE FROM N.A.
RP      MEDLINE=88121677; PubMed=3323826;
RA      Uchimura T., Lau P.C.K.;
RT      "Nucleotide sequences from the colicin E8 operon: homology with
RL      plasmid ColE2-P9".
RM      Mol. Gen. Genet. 209:489-493(1987).
RN      [2]
SEQUENCE FROM N.A.
RP      MEDLINE=88257046; PubMed=3290201;
RA      Toba M., Masaki H., Ohta T.;
RT      "Colicin E8, a DNase which indicates an evolutionary relationship
RL      between colicins E2 and E3".
RN      J. Bacteriol. 170:3237-3242(1988).
CC      [-] FUNCTION: THIS PLASMID-CODED BACTERICIDAL PROTEIN IS AN
CC      ENDONUCLEASE ACTIVE ON BOTH SINGLE- AND DOUBLE-STRANDED DNA BUT
CC      WITH UNDEFINED SPECIFICITY.
CC      [-] FUNCTION: COLICINS ARE POLYPEPTIDE TOXINS PRODUCED BY AND ACTIVE
CC      AGAINST; ESCHERICHIA COLI AND CLOSELY RELATED BACTERIA.
CC      [-] SIMILARITY: BELONGS TO THE NUCLEASE FAMILY OF COLICINS AND
CC      PYOSINS.
-----
THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement
or send an email to license@isb-sib.ch).
-----
EMBL; M21404; AAA23073.1; -
DR      EMBL; X06119; CAA29491.1; -
PIR; A28184; NDECE8.
DR      HSSP; Q47112; 7CEI.
DR      InterPro; IPR002711; HNH.
DR      Pfam; PF01844; HNH; 1.
DR      SMART; SM00507; HNHC; 1.
DR      Antibiotic; Bacteriocin; Plasmid; Hydrolase; Endonuclease;
KW      Zinc; Metal-binding.
FT      NON_TER      1
FT      METAL      173 173      ZINC (BY SIMILARITY).
FT      METAL      198 198      ZINC (BY SIMILARITY).
FT      METAL      202 202      ZINC (BY SIMILARITY).
SQ      SEQUENCE      205 AA; 23198 MW; B3E292480A669155 CRC64;

Query Match          71.11%; Score 32; DB 1; Length 205;
Best Local Similarity 77.8%; Pred. No. 13;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```



RESULT 2  
 L112 CAEEL STANDARD; PRT; 1429 AA.  
 AC P14585;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Lin-12 protein precursor.  
 GN Lin-12 OR R107.8.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Bristol N2;  
 RX MEDLINE=88334747; PubMed=3419531;  
 RA Yochem J., Weston K., Greenwald I.;  
 RT "The Caenorhabditis elegans lin-12 gene encodes a transmembrane  
 RL protein with overall similarity to Drosophila Notch.";  
 RN Nature 335:547-550(1988).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Bristol N2;  
 RX MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fraser A.,  
 RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,  
 RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,  
 RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkneen R.,  
 RA Sims M., Smaldon N., Smith A., Smith M., Sonnhammer E., Staden R.,  
 RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,  
 RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,  
 RA Wohlschlag P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RL elegans.";  
 CC -1- FUNCTION: LIN-12 IS INVOLVED IN SEVERAL CELL FATES DECISIONS THAT  
 CC REQUIRES CELL-CELL INTERACTIONS. IT IS POSSIBLE THAT LIN-12  
 CC ENCODES A MEMBRANE-BOUND RECEPTOR FOR A SIGNAL THAT ENABLES  
 CC EXPRESSION OF THE VENTRAL UTERINE PRECURSOR CELL FATE.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: HIGH, TO C.ELEGANS GLP-1.  
 CC -1- SIMILARITY: Contains 13 EGF-like domains.  
 CC -1- SIMILARITY: Contains 3 Lin/Notch repeats.  
 CC -1- SIMILARITY: Contains 5 ANK repeats.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M12069; AAA70191.1; -;  
 DR EMBL: Z14092; CAA78474.1; -;  
 DR PIR: S06434; S06434.  
 DR HSP: P00740; IEDM.  
 DR WormPep: R107.8; CE00274.  
 DR InterPro: IPR002110; ANK.  
 DR InterPro: IPR000152; Asx\_hydroxyl.  
 DR InterPro: IPR000742; EGF\_2.  
 DR InterPro: IPR001881; EGF\_Ca.  
 DR InterPro: IPR006209; EGF\_like.  
 DR InterPro: IPR002049; Laminin\_EGF.  
 DR InterPro: IPR000800; Notch.  
 DR Pfam: PF00023; ank; 6.  
 DR Pfam: PF00008; EGF; 13.

DR Pfam: PF00066; notch; 3.  
 DR PRINTS; PRO0011; EGF\_LAMININ.  
 DR PRINTS; PRO1452; NOTCH.  
 DR SMART; SM00248; ANK; 6.  
 DR SMART; SM00179; EGF\_CA; 2.  
 DR SMART; SM00004; NL; 2.  
 DR PROSITE; PSS0088; ANK\_REPEAT; 3.  
 DR PROSITE; PSS0297; ANK\_REPEAT\_REGION; 1.  
 DR PROSITE; PSS0010; ASX\_HYDROXYL; 3.  
 DR PROSITE; PSS0022; EGF\_1; 12.  
 DR PROSITE; PSS0186; EGF\_2; 11.  
 DR PROSITE; PSS0187; EGF\_CA; 2.  
 KW Differentiation; Repeat; ANK repeat; EGF-like domain; Transmembrane;  
 KW Glycoprotein; Signal.  
 FT SIGNAL 1 15 POTENTIAL.  
 FT CHAIN 16 1429 LIN-12 PROTEIN.  
 FT DOMAIN 16 908 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 909 931 POTENTIAL.  
 FT DOMAIN 932 1429 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 20 61 EGF-LIKE 1.  
 FT DOMAIN 114 150 EGF-LIKE 2.  
 FT DOMAIN 152 190 EGF-LIKE 3.  
 FT DOMAIN 201 246 EGF-LIKE 4.  
 FT DOMAIN 250 285 EGF-LIKE 5.  
 FT DOMAIN 287 323 EGF-LIKE 6.  
 FT DOMAIN 323 363 EGF-LIKE 7.  
 FT DOMAIN 365 402 EGF-LIKE 8.  
 FT DOMAIN 404 441 EGF-LIKE 9.  
 FT DOMAIN 449 492 EGF-LIKE 10.  
 FT DOMAIN 503 541 EGF-LIKE 11.  
 FT DOMAIN 543 579 EGF-LIKE 12.  
 FT DOMAIN 582 619 EGF-LIKE 13.  
 FT REPEAT 635 669 LIN/NOTCH 1.  
 FT REPEAT 670 710 LIN/NOTCH 2.  
 FT REPEAT 711 750 LIN/NOTCH 3.  
 FT REPEAT 1093 1122 ANK 1.  
 FT REPEAT 1126 1158 ANK 2.  
 FT REPEAT 1162 1194 ANK 3.  
 FT REPEAT 1206 1236 ANK 4.  
 FT REPEAT 1240 1269 ANK 5.  
 FT DISULFID 24 35 BY SIMILARITY.  
 FT DISULFID 29 49 BY SIMILARITY.  
 FT DISULFID 51 60 BY SIMILARITY.  
 FT DISULFID 118 129 BY SIMILARITY.  
 FT DISULFID 123 138 BY SIMILARITY.  
 FT DISULFID 140 149 BY SIMILARITY.  
 FT DISULFID 156 169 BY SIMILARITY.  
 FT DISULFID 163 178 BY SIMILARITY.  
 FT DISULFID 180 189 BY SIMILARITY.  
 FT DISULFID 205 227 BY SIMILARITY.  
 FT DISULFID 221 234 BY SIMILARITY.  
 FT DISULFID 236 245 BY SIMILARITY.  
 FT DISULFID 254 264 BY SIMILARITY.  
 FT DISULFID 259 273 BY SIMILARITY.  
 FT DISULFID 275 284 BY SIMILARITY.  
 FT DISULFID 291 302 BY SIMILARITY.  
 FT DISULFID 296 311 BY SIMILARITY.  
 FT DISULFID 313 322 BY SIMILARITY.  
 FT DISULFID 327 339 BY SIMILARITY.  
 FT DISULFID 334 351 BY SIMILARITY.  
 FT DISULFID 353 362 BY SIMILARITY.  
 FT DISULFID 369 381 BY SIMILARITY.  
 FT DISULFID 375 390 BY SIMILARITY.  
 FT DISULFID 392 401 BY SIMILARITY.  
 FT DISULFID 408 419 BY SIMILARITY.  
 FT DISULFID 413 429 BY SIMILARITY.  
 FT DISULFID 431 440 BY SIMILARITY.  
 FT DISULFID 507 518 BY SIMILARITY.  
 FT DISULFID 512 529 BY SIMILARITY.  
 FT DISULFID 531 540 BY SIMILARITY.  
 FT DISULFID 547 558 BY SIMILARITY.  
 FT DISULFID 552 567 BY SIMILARITY.  
 FT DISULFID 569 578 BY SIMILARITY.

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	33	73.3	482	1	MURC_PASMU	P57818	pasteurella
2	33	73.3	1429	1	L112_CAEEL	P14585	caenorhabditis
3	32	71.1	190	1	MOBA_RHOCA	Q9x7k0	rhodobacter
4	32	71.1	205	1	CEA8_ECOLI	PQ9882	escherichia
5	32	71.1	309	1	FTSQ_RHIMI	O30993	rhizobium m
6	32	71.1	427	1	DH12_SHEEP	P50168	ovis aries
7	32	71.1	581	1	CEA2_ECOLI	P04419	escherichia
8	32	71.1	582	1	CEA9_ECOLI	PQ9883	escherichia
9	32	71.1	1158	1	KCH2_CANFA	Q9tsz3	canis famill
10	32	71.1	1159	1	KCH2_HUMAN	Q12809	homo sapien
11	32	71.1	1161	1	KCH2_RABIT	Q8wny2	oryctolagus
12	32	71.1	1162	1	KCH2_MOUSE	Q53219	mus musculus
13	32	71.1	1163	1	KCH2_RAT	O08962	rattus norv
14	31	68.9	208	1	ATPF_MYCGE	P47643	mycoplasma
15	31	68.9	343	1	DHAS_ARCFU	O28766	archaeoglob
16	31	68.9	354	1	F16Q_MOUSE	P70695	mus musculus
17	31	68.9	400	1	PROV_SALTY	P17328	salmonella
18	31	68.9	490	1	MURD_MYLEE	P57995	mycobacteri
19	31	68.9	493	1	LOM2_THEAC	Q9h1x6	thermoplasm
20	30	66.7	376	1	Y582_METAL	Q8pzb2	methanosarc
21	30	66.7	1018	1	YDJ1_ECOLI	P77748	escherichia
22	29	64.4	140	1	YF98_METTH	O27635	methanobact
23	29	64.4	258	1	CYSH_VIBVU	Q8cwk6	vibrio vuln
24	29	64.4	263	1	Y002_PSEPK	P31856	pseudomonas
25	29	64.4	267	1	MSX2_HUMAN	P35548	homo sapien
26	29	64.4	292	1	HCHA_STAAM	Q99w58	staphylococ
27	29	64.4	292	1	HCHA_STAAW	Q8nxy2	staphylococ
28	29	64.4	352	1	MURG_CHLTR	O84766	chlamydia t
29	29	64.4	400	1	CYH2_HUMAN	Q99418	homo sapien
30	29	64.4	400	1	CYH2_MOUSE	P97695	mus musculu
31	29	64.4	494	1	COBQ_MYCTU	O53677	mycobacteri
32	29	64.4	530	1	YK03_YEAST	P36119	saccharomyc
33	29	64.4	582	1	HS60_SCHPO	Q09864	schizosacch

C>Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 24-May-2001  
C:Accession: T50872  
R:Nagashima, K.V.; Igarashi, N.; Harada, J.; Nagashima, S.; Matsuura, K.; Shimada, K.  
Submitted to the EMBL Data Library, November 1999  
A:Description: Determination of Nucleotide Sequences of Rubrivivax gelatinosus Photosynth  
A:Reference number: Z25270  
A:Accession: T50872  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-277 <NAG>  
A:Cross-references: EMBL:AB034704; PIDN:BAA94025.1  
A:Experimental source: strain IL144  
C:Genetics:  
A:Note: ORF277  
C:Superfamily: spore germination protein C2; bioc homology

Query Match 68.9%; Score 31; DB 2; Length 277;  
Best Local Similarity 60.0%; Pred. No. 69;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AMLAAQERRV 10  
||| : |||  
Db 105 ALMAAADRRV 114

RESULT 12  
A83299  
hypothetical protein PA2778 [imported] - Pseudomonas aeruginosa (strain PA01)  
C:Species: Pseudomonas aeruginosa  
C>Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: A83299  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
A:Reference number: A82950; MUID:20437337; PMID:10984043  
A:Accession: A83299  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-293 <STO>  
A:Cross-references: GB:AE004705; GB:AE004091; NID:99948851; PIDN:AAG06166.1; GSPDB:GN001  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA2778

Query Match 68.9%; Score 31; DB 2; Length 293;  
Best Local Similarity 70.0%; Pred. No. 73;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AMLAAQERRV 10  
||| : |||  
Db 218 AMLTSVERRV 227

RESULT 13  
A69438  
aspartate-semialdehyde dehydrogenase (asd) homolog - Archaeoglobus fulgidus  
C:Species: Archaeoglobus fulgidus  
C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 04-Nov-2002  
C:Accession: A69438  
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.  
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
Nature 390, 364-370, 1997  
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.  
Smith, H.O.; Woese, C.R.; Venter, J.C.  
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo  
A:Reference number: A69250; MUID:98049343; PMID:9389475  
A:Accession: A69438  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-343 <KLE>

A:Cross-references: GB:AE000998; GB:AE000782; NID:g2689321; PIDN:AAB89738.1; PID:g264  
C:Superfamily: aspartate-semialdehyde dehydrogenase

Query Match 68.9%; Score 31; DB 2; Length 343;  
Best Local Similarity 87.5%; Pred. No. 85;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LAAQERRV 10  
||| ||||  
Db 33 LAASERRV 40

RESULT 14  
S46245  
RAE-30 protein - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 10-Dec-1994 #sequence\_revision 12-May-1995 #text\_change 20-Jun-2000  
C:Accession: S46245  
R:Nomura, M.; Takihara, Y.; Yasunaga, T.; Shimada, K.  
FEBS Lett. 348, 201-205, 1994  
A:Title: One of the retinoic acid-inducible cDNA clones in mouse embryonal carcinoma  
A:Reference number: S46245; MUID:94307425; PMID:8034042  
A:Accession: S46245  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-354 <NOM>  
A:Cross-references: GB:D42083; NID:g575935; PIDN:BAA07678.1; PID:g575936  
C:Superfamily: fructose-bisphosphatase

Query Match 68.9%; Score 31; DB 2; Length 354;  
Best Local Similarity 66.7%; Pred. No. 88;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AMLAAQERR 9  
||| : ||||  
Db 103 AVITAQERR 111

RESULT 15  
S62692  
L-amino acid oxidase - Synecococcus sp. (PCC 6301)  
C:Species: Synecococcus sp.  
A:Variety: PCC 6301  
C>Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 29-Aug-1997  
C:Accession: S62692; S62687  
R:Boekholt, R.; Masepohl, B.; Kruft, V.; Wittmann-Liebold, B.; Pistorius, E.K.  
Biochim. Biophys. Acta 1264, 289-293, 1995  
A:Title: Partial amino acid sequence of an L-amino acid oxidase from the cyanobacteri  
A:Reference number: S62687; MUID:96138550; PMID:8547316  
A:Accession: S62692  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-355 <BOC>  
A:Cross-references: EMBL:Z48565  
A:Accession: S62687  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 52-68 <BOC>

Query Match 68.9%; Score 31; DB 2; Length 355;  
Best Local Similarity 87.5%; Pred. No. 88;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LAAQERRV 10  
||| ||||  
Db 95 LAFQERRV 102

Search completed: October 7, 2003, 13:32:40  
Job time : 7.63291 secs

Query Match 71.1%; Score 32; DB 2; Length 427;  
Best Local Similarity 77.8%; Pred. No. 65;  
Matches 7; Conservative 1; Mismatches 0; Gaps 0;

QY 1 AMLAAQERR 9  
1:|||||  
328 ALLAAQPR 336

RESULT 7  
NDBCE2  
colicin E2 (EC 3.1.21.-) - Escherichia coli plasmid Cole2-P9  
N:Alternate names: colicin E2 activity protein  
C:Species: Escherichia coli  
C:Date: 17-Mar-1987 #sequence\_revision 09-May-1997 #text\_change 11-Jun-1999  
C:Accession: I40687; A00789  
R:Cole, S.T.; Saint-Joanis, B.; Pugsley, A.P.  
Mol. Gen. Genet. 198, 465-472, 1985  
A:Title: Molecular characterisation of the colicin E2 operon and identification of its P  
A:Reference number: I40687; MUID:85239907; PMID:3892228  
A:Accession: I40687  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-581 <RES>  
A:Cross-references: GB:M29885; NID:g144366; PIDN:AAA23068.1; PID:g144367  
A:Experimental source: plasmid Cole2  
R:Lau, P.C.K.; Rowsome, R.W.; Zuker, M.; Visentin, L.P.  
Nucleic Acids Res. 12, 8733-8745, 1984  
A:Title: Comparative nucleotide sequences encoding the immunity proteins and the carboxy  
A:Reference number: A00789; MUID:85062845; PMID:6095211  
A:Accession: A00789  
A:Molecule type: DNA  
A:Residues: 377-428, 'RS', 432-472, 'R', 474-503, 'L', 505-581 <LAU>  
A:Cross-references: GB:X01163; NID:g41800; PIDN:CAA25609.1; PID:g809683  
A:Experimental source: plasmid Cole2  
C:Comment: This plasmid-coded bactericidal protein is an endonuclease active on both sin  
C:Genetics:  
A:Gene: ceaB; col  
A:Genome: plasmid  
C:Superfamily: cloacin DFI3 protein  
C:Keywords: bacteriocin; endonuclease; hydrolase; toxin

Query Match 71.1%; Score 32; DB 1; Length 581;  
Best Local Similarity 77.8%; Pred. No. 89;  
Matches 7; Conservative 1; Mismatches 0; Gaps 0;

QY 1 AMLAAQERR 9  
1:|||||  
424 ALSAAQERR 432

RESULT 8  
I38465  
probable potassium channel subunit - human  
C:Species: Homo sapiens (man)  
C:Date: 31-May-1996 #sequence\_revision 31-May-1996 #text\_change 05-Nov-1999  
C:Accession: I38465  
R:Warme, J.W.; Ganetzky, B.  
Proc. Natl. Acad. Sci. U.S.A. 91, 3438-3442, 1994  
A:Title: A family of potassium channel genes related to eag in Drosophila and mammals.  
A:Reference number: A54953; MUID:94211879; PMID:8159766  
A:Accession: I38465  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1159 <RES>  
A:Cross-references: EMBL:U04270; NID:g487737; PIDN:AAA62473.1; PID:g487738  
C:Superfamily: cAMP receptor protein cyclic nucleotide-binding domain homology  
F,742-858/Domain: cAMP receptor protein cyclic nucleotide-binding domain homology <CAP>

Query Match 71.1%; Score 32; DB 2; Length 1159;  
Best Local Similarity 60.0%; Pred. No. 1.8e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 AMLAAQERR 10  
1:|||||  
85 ALLGAERKV 94

RESULT 9  
F64244  
ATP synthase B chain (atpF) - Mycoplasma genitalium  
C:Species: Mycoplasma genitalium  
C:Date: 17-Nov-1995 #sequence\_revision 17-Nov-1995 #text\_change 07-Dec-1999  
C:Accession: F64244  
R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.  
M.; Furmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick,  
C.A.; Venter, J.C.  
Science 270, 397-403, 1995  
A:Title: The minimal gene complement of Mycoplasma genitalium.  
A:Reference number: A64200; MUID:96026346; PMID:7569993  
A:Accession: F64244  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-208 <TIGR>  
A:Cross-references: GB:U39726; GB:L43967; NID:g1046113; PID:g1046117; TIGR:MG403  
A:Experimental source: strain G-37  
C:Genetics:  
A:Genetic code: SGC3  
A:Superfamily: H+-transporting ATP synthase chain B  
C:Keywords: ATP biosynthesis

Query Match 68.9%; Score 31; DB 2; Length 208;  
Best Local Similarity 100.0%; Pred. No. 52;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AMLAAQ 7  
1:|||||  
172 AMLAAQ 178

RESULT 10  
E70822  
hypothetical protein RV0730 - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
C:Accession: E70822  
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon  
J.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd,  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno  
A:Reference number: A70500; MUID:98295987; PMID:9634230  
A:Accession: E70822  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-242 <COL>  
A:Cross-references: GB:AL021958; GB:AL123456; NID:g3261536; PIDN:CAA17497.1; PID:g291  
A:Experimental source: strain H37Rv  
C:Genetics:  
A:Gene: RV0730

Query Match 68.9%; Score 31; DB 2; Length 242;  
Best Local Similarity 77.8%; Pred. No. 60;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AMLAAQERR 9  
1:|||||  
30 ALLAALERR 38

RESULT 11  
T50872  
hypothetical protein ORF277 [imported] - Rubrivivax gelatinosus  
C:Species: Rubrivivax gelatinosus

C;Superfamily: unassigned ankryrin repeat proteins; ankryrin repeat homology; EGF homology  
C;Keywords: glycoprotein; transmembrane protein  
F;254-284/Domain: EGF homology <EGF1>  
F;507-540/Domain: EGF homology <EGF>  
F;547-578/Domain: EGF homology <EGF2>  
F;909-931/Domain: transmembrane status predicted <TM>  
F;1093-1125/Domain: ankryrin repeat homology <AN1>  
F;1206-1238/Domain: ankryrin repeat homology <AN2>  
F;1240-1272/Domain: ankryrin repeat homology <AN3>  
Query Match 73.3%; Score 33; DB 2; Length 1429;  
Best Local Similarity 77.8%; Pred. No. 1.3e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 2 MLAAQERRV 10  
Db 1246 MLAAQEGRI 1254  
||||| |  
RESULT 3  
NDCE8  
colicin E8 (EC 3.1.21.-) - Escherichia coli plasmid Cole8-J (fragment)  
N;Alternate names: colicin E8 complex, protein A  
C;Species: Escherichia coli  
C;Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 11-Jun-1999  
C;Accession: A28184; S01080  
R;Toba, M.; Masaki, H.; Ohta, T.  
J. Bacteriol. 170, 3237-3242, 1988  
A;Title: Colicin E8, a DNase which indicates an evolutionary relationship between colicins  
A;Reference number: A91874; MUID:88257046; PMID:3290201  
A;Accession: A28184  
A;Molecule type: DNA  
A;Residues: 1-205 <TO>  
A;Cross-references: GB:M21404; NID:g144380; PIDN:AAA23073.1; PID:g144381  
R;Ochimura, T.; Lau, P.C.K.  
Mol. Gen. Genet. 209, 489-493, 1987  
A;Title: Nucleotide sequences from the colicin E8 operon: homology with plasmid ColE2-P9  
A;Reference number: S01080; MUID:88121677; PMID:3323826  
A;Accession: S01080  
A;Molecule type: DNA  
A;Residues: 1-205 <UCH>  
A;Cross-references: EMBL:X06119; NID:g41136; PIDN:CAA29491.1; PID:g809672  
A;Note: the authors translated the codon AAA for residue 75 as Ala  
C;Genetics:  
A;Gene: col  
A;Genome: plasmid  
C;Superfamily: cloacin DF13 protein  
C;Keywords: bacteriocin; endonuclease; hydrolase; toxin  
Query Match 71.1%; Score 32; DB 1; Length 205;  
Best Local Similarity 77.8%; Pred. No. 31;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 AMLAQERR 9  
Db 48 ALSAQERR 56  
||||| |  
RESULT 4  
PQ0032  
colicin E9 (EC 3.1.21.-) - Escherichia coli plasmid Cole9-J (fragment)  
C;Species: Escherichia coli  
C;Date: 07-Sep-1990 #sequence\_revision 22-Nov-1996 #text\_change 11-Jun-1999  
C;Accession: PQ0032; S03087  
R;Lau, P.C.K.; Condie, J.A.  
Mol. Gen. Genet. 217, 269-277, 1989  
A;Title: Nucleotide sequences from the colicin E5, E6 and E9 operons: presence of a degenerate  
A;Reference number: JQ0326; MUID:89364708; PMID:2549375  
A;Accession: PQ0032  
A;Molecule type: DNA  
A;Residues: 1-205 <LAU>  
A;Cross-references: GB:X15858; NID:g40548; PIDN:CAA33862.1; PID:g809666  
A;Experimental source: strain W3110

R;Eaton, T.; James, R.  
Nucleic Acids Res. 17, 1761, 1989  
A;Title: Complete nucleotide sequence of the colicin E9 (cei) gene.  
A;Reference number: S03087; MUID:89160332; PMID:2646600  
A;Accession: S03087  
A;Status: translation not shown  
A;Molecule type: DNA  
A;Residues: 123-124, 'GRLYGR', 132-133, 'R', 135-201, 'L', 203-205 <EAT>  
A;Cross-references: EMBL:X12591  
A;Note: Colicin E9 has desoxyribonuclease activity  
C;Genetics:  
A;Gene: colE9; cei  
A;Genome: plasmid  
C;Superfamily: cloacin DF13 protein  
C;Keywords: bacteriocin; endonuclease; hydrolase; toxin  
Query Match 71.1%; Score 32; DB 1; Length 205;  
Best Local Similarity 77.8%; Pred. No. 31;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 AMLAQERR 9  
Db 48 ALSAQERR 56  
||||| |  
RESULT 5  
G87468  
ubiquinol oxidase subunit III [imported] - Caulobacter crescentus  
C;Species: Caulobacter crescentus  
C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 10-May-2001  
C;Accession: G87468  
R;Nierman, W.C.; Feidblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A;Title: Complete Genome Sequence of Caulobacter crescentus.  
A;Reference number: A87249; MUID:21173698; PMID:11259647  
A;Accession: G87468  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-208 <STO>  
A;Cross-references: GB:AE005673; NID:g13423197; PIDN:AAK23747.1; GSPDB:GN00148  
C;Genetics:  
A;Gene: CCI771  
C;Superfamily: cytochrome-c oxidase chain III  
Query Match 71.1%; Score 32; DB 2; Length 208;  
Best Local Similarity 70.0%; Pred. No. 32;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 1 AMLAQERRV 10  
Db 89 AMTAAQAREV 98  
||||| |  
RESULT 6  
A55353  
11beta-hydroxysteroid dehydrogenase (EC 1.1.1.146), isoform K - sheep  
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C;Date: 06-Feb-1995 #sequence\_revision 06-Feb-1995 #text\_change 17-Mar-1999  
C;Accession: A55353  
R;Agarwal, A.K.; Mune, T.; Monder, C.; White, P.C.  
J. Biol. Chem. 269, 25959-25962, 1994  
A;Title: NAD(+)-dependent isoform of 11beta-hydroxysteroid dehydrogenase. Cloning and  
A;Reference number: A55353; MUID:95014413; PMID:7929304  
A;Accession: A55353  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-427 <AGA>  
A;Cross-references: GB:U14128  
C;Superfamily: retinol dehydrogenase; short-chain alcohol dehydrogenase homology  
C;Keywords: kidney; oxidoreductase  
F;83-265/Domain: short-chain alcohol dehydrogenase homology <SADH>

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:19:12 ; Search time 5.63291 Seconds  
(without alignments)  
170.726 Million cell updates/sec

Title: US-09-807-512-25  
Perfect score: 45  
Sequence: 1 AMLAAQERRV 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues  
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_76:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	33	73.3	306	D87531	glycosyl transferase
2	33	73.3	1429	S06434	homeotic protein 1
3	32	71.1	205	NDECER	colicin E8 (EC 3.1
4	32	71.1	205	PQ0032	colicin E9 (EC 3.1
5	32	71.1	208	G87468	ubiquinol oxidase
6	32	71.1	427	A35353	libeta-hydroxyster
7	32	71.1	581	NDECER	colicin E2 (EC 3.1
8	32	71.1	1159	I38465	probable potassium
9	31	68.9	208	F64244	ATP synthase B cha
10	31	68.9	242	E70822	hypothetical prote
11	31	68.9	277	T50872	hypothetical prote
12	31	68.9	293	A83299	hypothetical prote
13	31	68.9	343	A69438	aspartate-semialde
14	31	68.9	354	S46245	RAE-30 protein - m
15	31	68.9	355	S62692	L-amino acid oxida
16	31	68.9	383	E87680	hypothetical prote
17	31	68.9	400	QREBVT	glycine betaine/pr
18	31	68.9	401	G84160	glucose-1-phosphat
19	31	68.9	490	B87023	UDP-N-acetylmutram
20	31	68.9	500	A87645	efflux system prot
21	31	68.9	1047	T25782	hypothetical prote
22	30	66.7	165	A71105	hypothetical prote
23	30	66.7	168	C75056	hypothetical prote
24	30	66.7	233	T35286	hypothetical prote
25	30	66.7	235	E75375	probable methyltra
26	30	66.7	291	C87700	conserved hypothet
27	30	66.7	295	E83058	UTP-glucose-1-phos
28	30	66.7	319	T36857	hypothetical prote
29	30	66.7	331	AF3544	conserved hypothet
					hypothetical membr

30	30	66.7	394	2	A70842	probable amiB prot
31	30	66.7	550	2	H95906	probable protein 1
32	30	66.7	550	2	H95924	hypothetical prote
33	30	66.7	583	2	T49359	hypothetical prote
34	30	66.7	733	2	C87655	penicillin-binding
35	30	66.7	995	2	T32466	hypothetical prote
36	30	66.7	1018	2	B90928	probable oxidase I
37	30	66.7	1018	2	F85776	probable oxidase y
38	30	66.7	1018	2	G64926	probable iron-sulf
39	29	64.4	72	2	JS0654	hypothetical 8.2K
40	29	64.4	118	2	T34991	hypothetical prote
41	29	64.4	130	2	B95873	hypothetical prote
42	29	64.4	140	2	C69080	conserved hypothet
43	29	64.4	168	2	B83975	hypothetical prote
44	29	64.4	187	2	T34834	probable transfera
45	29	64.4	197	2	H75563	transcription regu

ALIGNMENTS

RESULT 1

D87531  
glycosyl transferase family protein CC2277 [Imported] - Caulobacter crescentus  
C:Species: Caulobacter crescentus  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 10-May-2001  
C:Accession: D87531  
R:NIERMAN, W.C.; FELDBLYUM, T.V.; PAULSEN, I.T.; NELSON, K.E.; EISEN, J.; HEIDELBERG, B.; LAUB, M.T.; DEBOY, R.T.; DODSON, R.J.; DURKIN, A.S.; GWINN, M.L.; HAFT, D.H.; KO N, J.; ERMOLAEVA, M.; WHITE, O.; SALZBERG, S.L.; SHAPIRO, L.; VENTER, J.C.; FRASER, C. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of Caulobacter crescentus.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: D87531  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-306 <STO>  
A:Cross-references: GB:AE005673; NID:gl3423792; PIDN:AAK24248.1; GSPDB:GN00148  
C:Genetics:  
A:Gene: CC2277  
C:Superfamily: Rhizobium succinoglycan biosynthesis glycosyltransferase

Query Match 73.3%; Score 33; DB 2; Length 306;  
Best Local Similarity 87.5%; Pred. No. 29;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AMLAAQER 8  
I:|||||  
DB 105 ALLAAQER 112

RESULT 2

S06434

homeotic protein lin-12 precursor - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 29-Jan-1993 #sequence\_revision 29-Jan-1993 #text\_change 20-Sep-1999  
C:Accession: S06434; A24769

R:YOCHEN, J.; WESTON, K.; GREENWALD, I. Nature 335, 547-550, 1988  
A:Title: The Caenorhabditis elegans lin-12 gene encodes a transmembrane protein with  
A:Reference number: S06434; MUID:88334747; PMID:3419531  
A:Accession: S06434  
A:Molecule type: DNA  
A:Residues: 1-1429 <YOC>  
A:Cross-references: EMBL:M12069; NID:gl56357; PIDN:AAA70191.1; PID:gl56358  
R:Greenwald, I. Cell 43, 583-590, 1985  
A:Reference number: A24769; MUID:86079540; PMID:3000611  
A:Accession: A24769  
A:Molecule type: DNA  
A:Residues: 173-712 <GRE>  
C:Genetics:  
A:Introns: 50/2; 90/1; 109/1; 172/3; 545/1; 589/2; 632/2; 1273/3; 1389/3

```

XX SQ Sequence 11 AA;
Query Match 91.1%; Score 41; DB 20; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MAAQERRV 10
Db 1 MAAQERRV 9
|||||

RESULT 14
AAY05982
ID AAY05982 standard; Peptide; 10 AA.
XX AC
XX AAY05982;
XX 16-AUG-1999 (first entry)
XX DE
XX Human cancer antigen NY ESO-1/CAG-3 ORF2 cancer peptide.
XX NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;
XX leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;
XX metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;
XX uterine cancer; breast cancer; prostate cancer; ovarian cancer;
XX cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;
XX liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;
XX vaccine.
XX OS
XX Homo sapiens.
XX WO9918206-A2.
XX 15-APR-1999.
XX 21-SEP-1998; 98WO-US19609.
XX 08-OCT-1997; 97US-0061428.
XX (USSS ) US DEPT HEALTH & HUMAN SERVICES.
XX Rosenberg SA, Wang RF;
XX WPI; 1999-277270/23.
XX N-PSDB; AAX58601.
XX Cancer antigen NY ESO1/CAG-3
XX Claim 26; Page 65; 88pp; English.
XX The present sequence represents a cancer peptide that is based on
XX amino acid residues 19-27 of human ESO-1/CAG-3 (or CAG-3) ORF2
XX (see AAY05966), a new and potent tumour antigen capable of eliciting
XX an antigen specific immune response by T cells. Cancer peptides
XX derived from CAG-3 ORF2, CAG-3 ORF1 (see AAY05965), portions of them
XX and their variants (see AAY05967-87), are useful as cancer vaccines
XX that protect against cancer. The invention provides: vectors and
XX host cells (also useful as vaccines); a method of diagnosis of
XX cancer or precancer; a transgenic animal; antisense oligonucleotides
XX that inhibit expression of the cancer peptide or tumour antigen;
XX antibodies reacting with a CAG-3 cancer peptide, useful in
XX diagnostic and detection assays; and methods for preventing or
XX inhibiting cancer by administering a cancer peptide, with or without
XX an HLA molecule. The cancer peptides form part of, or are derived
XX from, cancers such as primary or metastatic melanoma, thymoma,
XX lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine
XX cancer, cervical cancer, bladder cancer, kidney cancer and
XX adenocarcinomas such as breast, prostate, ovarian, pancreatic and
XX thyroid cancers. Melanoma is treated by inducing cancer-specific T
XX cells in vitro for subsequent return to a patient.
XX Sequence 10 AA;

```

```

Query Match 80.0%; Score 36; DB 20; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAAQERRV 10
Db 1 LAAQERRV 8
|||||

RESULT 15
AAB31332
ID AAB31332 standard; peptide; 10 AA.
XX AC
XX AAB31332;
XX 20-APR-2001 (first entry)
XX DE
XX Exemplary antigen characteristic of tumours and derived from NY-ESO-1.
XX MAGE-A1; HLA; human leukocyte antigen; CD4+ T lymphocyte; cancer;
XX MAGE-A1 HLA class II-binding protein; vaccine.
XX OS
XX Homo sapiens.
XX WO200078806-A1.
XX 28-DEC-2000.
XX 14-JUN-2000; 2000WO-US16287.
XX 18-JUN-1999; 99US-0336091.
XX (LUDW-) LUDWIG INST CANCER RES.
XX Van Snick J, Lethe B, Chaux P, Boon-Falleur T, Van Der Bruggen P;
XX WPI; 2001-102698/11.
XX Novel MAGE-A1 human leukocyte antigen class II peptides which bind to
XX and are presented to the class II molecules, useful for inducing immune
XX response and treating cancers characterized by expression of MAGE-A1 -
XX Disclosure; Page 32; 78pp; English.
XX AAB31302-59 represent exemplary antigens which are characteristic
XX of tumours. They can be used to enhance the immune response of vaccines
XX comprising peptides derived from human MAGE-A1 HLA (human leukocyte
XX antigen) class II-binding protein. Peptides derived from the MAGE-A1 HLA
XX binding protein stimulate the activity and proliferation of CD4+ T
XX lymphocytes. The MAGE-A1 HLA binding protein is useful as a diagnostic
XX agent for diagnosing a disorder characterized by expression of MAGE-A1.
XX The protein is used for treating a disorder characterized by expression
XX of MAGE-A1 such as cancers e.g. melanoma, squamous cell carcinomas,
XX colorectal carcinomas, osteosarcomas, and lymphocytic leukaemias.
XX Peptides derived from the MAGE-A1 HLA binding protein are useful in the
XX production of anti-tumour vaccines.
XX Sequence 10 AA;

Query Match 80.0%; Score 36; DB 22; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAAQERRV 10
Db 1 LAAQERRV 8
|||||

Search completed: October 7, 2003, 13:25:33
Job time : 18.5823 secs

```

CC useful for modulating immune responses preferably directed against a  
 CC pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix,  
 CC colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone  
 CC liver, oesophagus, brain, testicle, uterus), as potentiating agents.  
 CC Compositions comprising the polypeptide may be used in the treatment or  
 CC prophylaxis against viral (such as infections caused by HIV (human  
 CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis  
 CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial  
 CC (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,  
 CC Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic  
 CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,  
 CC Trypanosoma, Toxoplasma and Giardia) infections. The present  
 CC sequence is a vaccine protein of the invention.

XX SQ Sequence 3541 AA;  
 Query Match 100.0%; Score 45; DB 23; Length 3541;  
 Best Local Similarity 100.0%; Pred. No. 9.4; 0; Gaps 0;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0;  
 OY 1 AMLAQERRV 10  
 DB 183 AMLAQERRV 192  
 |||||

RESULT 12  
 AAY70859  
 ID AAY70859 standard; peptide; 9 AA.

XX AC AAY70859;

XX DT 31-JUL-2000 (first entry)

XX DE CAMEL17 immunogenic peptide of human CAMEL protein.

XX KW CAMEL; CTL-recognised Antigen on MELANOMA; cytotoxic T lymphocyte; CTL;  
 KW tumour-associated antigen; LAGE-1; NY-ESO-1; anticancer; melanoma;  
 KW human; cancer; immunotherapy; immunogenic peptide; immune response.

XX OS Homo sapiens.

XX PN WO200023584-A1.

XX PD 27-APR-2000.

XX PF 15-OCT-1999; 99WO-EF07832.

XX PR 16-OCT-1998; 98EP-0119583.

XX PA (BOEH ) BOEHRINGER INGELHEIM INT GMBH.  
 PA (UYHO-) UNIV HOSPITAL LEIDEN.

XX PI Schrier PI, Aarnoudse CA, Heider K, Klade C;

XX PS WPI; 2000-339685/29.

XX PT Tumor-associated antigen useful for cancer immunotherapy is encoded by  
 PT the open reading frame of LAGE-1 (a tumor-specific antigen) CDNA -

XX PS Claim 8; Page 34; 73pp; English.

XX CC The present sequence is an immunogenic peptide CAMEL 17, of the human  
 CC tumour-associated antigen CAMEL (Cytotoxic T lymphocytes (CTL)-recognised  
 CC Antigen on MELANOMA). This peptide has the potential to bind to HLA-A2  
 CC and corresponds to residues 17-25 of the CAMEL protein. The CAMEL protein  
 CC is encoded by the LAGE-1 gene, a tumour-specific antigen. It is different  
 CC from the LAGE-1 protein, since it is translated from a different open  
 CC reading frame (ORF-1). It shows strong homology with NY-ESO-1, a melanoma  
 CC specific tumour antigen. The tumour-associated antigen displayed on  
 CC melanoma cells is recognised by cytotoxic T lymphocytes. CAMEL is  
 CC expressed in tumour cell lines, tumour tissues (e.g. breast and lung) and  
 CC in restricted number of healthy tissues. This sequence has anticancer  
 CC activity. CAMEL tumour antigen and immunogenic peptides derived from it

CC are useful for cancer immunotherapy. They have the potential to induce an  
 CC immune response, by eliciting a CTL response. The DNA molecule is used to  
 CC construct recombinant or fusion proteins.

XX SQ Sequence 9 AA;

Query Match 91.1%; Score 41; DB 21; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 MLAQERRV 10  
 |||||

DB 1 MLAQERRV 9

RESULT 13

AAY05983

ID AAY05983 standard; Peptide; 11 AA.

XX AC AAY05983;

XX DT 16-AUG-1999 (first entry)

XX DE Human cancer antigen NY ESO-1/CAG-3 ORF2 cancer peptide.

XX KW NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;  
 KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
 KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
 KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
 KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
 KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
 KW vaccine.

XX OS Homo sapiens.

XX PN WO9918206-A2.

XX PD 15-APR-1999.

XX PF 21-SEP-1998; 98WO-US19609.

XX PR 08-OCT-1997; 97US-0061428.

XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX PI Rosenberg SA, Wang RF;

XX DR WPI; 1999-277270/23.

XX PT Cancer antigen NY ESO1/CAG-3

XX PS Disclosure; Page 12; 88pp; English.

XX CC The present sequence represents a cancer peptide that corresponds  
 CC to amino acid residues 17-27 of human ESO-1/CAG-3 (or CAG-3) ORF2  
 CC (see AAY05966), a new and potent tumour antigen capable of eliciting  
 CC an antigen specific immune response by T cells. Cancer peptides  
 CC derived from CAG-3 ORF2, CAG-3 ORF1 (see AAY05965), portions of them  
 CC and their variants (see AAY05967-87), are useful as cancer vaccines  
 CC that protect against cancer. The invention provides: vectors and  
 CC host cells (also useful as vaccines); a method of diagnosis of  
 CC cancer or precancer; a transgenic animal; antisense oligonucleotides  
 CC that inhibit expression of the cancer peptide or tumour antigen;  
 CC antibodies reacting with a CAG-3 cancer peptide, useful in  
 CC diagnostic and detection assays; and methods for preventing or  
 CC inhibiting cancer by administering a cancer peptide, with or without  
 CC an HLA molecule. The cancer peptides form part of, or are derived  
 CC from, cancers such as primary or metastatic melanoma, thymoma,  
 CC lymphoma, sarcoma, lung cancer, liver cancer, kidney cancer and  
 CC cancer, cervical cancer, bladder cancer, kidney cancer and  
 CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and  
 CC thyroid cancers. Melanoma is treated by inducing cancer-specific T  
 CC cells in vitro for subsequent return to a patient.



```

Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AMLAAQERRY 10
   |||||
Db 16 AMLAAQERRY 25

RESULT 10
AAY70854
ID AAY70854 standard; Protein; 109 AA.
XX
AC AAY70854;
XX
DT 31-JUL-2000 (first entry)
XX
DE Human CTL-recognised Antigen on MELANOMA (CAMEL) protein.
XX
KW CAMEL; CTL-recognised Antigen on MELANOMA; cytotoxic T lymphocyte; CTL;
KW tumour-associated antigen; LAGE-1; NY-ESO-1; anticancer; melanoma;
KW cancer; immunotherapy; immunogenic peptide; immune response.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..11
FT /label= CTL_epitope
FT /note= "Immunogenic peptide with ability to elicit a
FT CTL response"
FT Peptide 2..10
FT /label= Immunogenic_peptide
FT /note= "Specific for HLA-A3"
FT Peptide 2..11
FT /label= CTL_epitope
FT /note= "Immunogenic peptide with ability to elicit a
FT CTL response"
FT Peptide 10..18
FT /label= CAMEL_10
FT /note= "Specific for HLA-A2"
FT Peptide 16..25
FT /label= CAMEL_16
FT /note= "Specific for HLA-A2"
FT Peptide 17..25
FT /label= CAMEL_17
FT /note= "Specific for HLA-A2"
FT Peptide 51..59
FT /label= Immunogenic_peptide
FT /note= "Specific for HLA-A3101"
FT Peptide 101..109
FT /label= Immunogenic_peptide
XX
PN WO200023584-A1.
XX
PD 27-APR-2000.
XX
PF 15-OCT-1999; 99WO-EP07832.
XX
PR 16-OCT-1998; 98EP-0119583.
XX
PA (BOEH ) BOEHRINGER INGELHEIM INT GMBH.
PA (UYHO-) UNIV HOSPITAL LEIDEN.
XX
PI Schrier PI, Aarnoudse CA, Heider K, Klade C;
XX
DR WPI; 2000-339685/29.
DR N-PSDB; AAD00149.
XX
PT Tumor-associated antigen useful for cancer immunotherapy is encoded by
PT the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA -
XX
XX Claim 1; Page 55; 73pp; English.
XX
XX The present protein sequence is the human tumour-associated antigen CAMEL

```

```

CC (Cytotoxic T lymphocytes (CTL)-recognised Antigen on MELANOMA). CAMEL
CC protein is encoded by the LAGE-1 gene, a tumour-specific antigen. It is
CC different from the LAGE-1 protein, since it is translated from a
CC different open reading frame (ORF-1). It shows strong homology with
CC NY-ESO-1, a melanoma specific tumour antigen. The tumour-associated
CC antigen displayed on melanoma cells is recognised by cytotoxic T
CC lymphocytes. CAMEL is expressed in tumour cell lines, tumour tissues
CC (e.g. breast and lung) and in restricted number of healthy tissues. This
CC sequence has anticancer activity. CAMEL tumour antigen and immunogenic
CC peptides derived from it are useful for cancer immunotherapy. They have
CC the potential to induce an immune response, by eliciting a CTL response.
CC The DNA molecule is used to construct recombinant or fusion proteins.
XX
SQ Sequence 109 AA;
   Query Match 100.0%; Score 45; DB 21; Length 109;
   Best Local Similarity 100.0%; Pred. No. 0.28;
   Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AMLAAQERRY 10
   |||||
Db 16 AMLAAQERRY 25

RESULT 11
AAU85130
ID AAU85130 standard; Protein; 3541 AA.
XX
AC AAU85130;
XX
DT 08-MAY-2002 (first entry)
XX
DE Human melanoma specific savine.
XX
KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;
KW viral infection; human immunodeficiency virus; melanoma;
KW bacterial infection; Salmonella; Legionella; parasitic infection;
KW Trypanosoma; Toxoplasma; Giardia.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200190197-A1.
XX
PD 29-NOV-2001.
XX
PF 25-MAY-2001; 2001WO-AU00622.
XX
PR 26-MAY-2000; 2000AU-0007761.
XX
PA (AUSU ) UNIV AUSTRALIAN NAT.
XX
PI Thomson SA, Ramshaw IA;
XX
DR WPI; 2002-147575/19.
DR N-PSDB; ABK36950.
XX
XX New synthetic polypeptides having several different segments of at
XX least one parent polypeptide linked together differently compared to
XX the linkage in the parent polypeptide, for inducing immune response
XX against a pathogen or cancer
XX
XX Example 3; Fig 27; 364pp; English.
XX
XX The invention relates to a new synthetic polypeptide (I) comprising
XX several different segments of at least one parent polypeptide linked
XX together in a different relationship relative to their linkage in the
XX parent polypeptide to impede, abrogate or otherwise alter at least one
XX function associated with the parent polypeptide and for inducing an
XX immune response against a pathogen or cancer. Also included are a
XX synthetic polynucleotide encoding and a computer system for
XX designing the synthetic polypeptides. The synthetic polypeptides and
XX polynucleotides are referred to as a Savine. The synthetic polypeptide is

```

CC primary or metastatic melanoma, thymoma, lymphoma, sarcoma, lung  
 CC cancer, liver cancer, leukemia, uterine cancer, cervical cancer,  
 CC bladder cancer, kidney cancer and adenocarcinomas such as breast,  
 CC prostate, ovarian, pancreatic and thyroid cancers. Melanoma is  
 CC treated by inducing cancer-specific T cells in vitro for subsequent  
 CC return to a patient.

XX Sequence 58 AA;

SQ Query Match 100.0%; Score 45; DB 20; Length 58;

Best Local Similarity 100.0%; Pred. No. 0.15;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AMLAAQERRV 10  
 DB 16 AMLAAQERRV 25

RESULT 8

AAY70863

ID AAY70863 standard; Protein; 58 AA.

XX AAY70863;

AC 31-JUL-2000 (first entry)

XX Human tumour antigen, NY-ESO-1 short variant protein.

XX NY-ESO-1; CAMEL; CTL-recognised Antigen on MELANOMA; human; cancer; CTL;  
 KW cytotoxic T lymphocyte; tumour-associated antigen; LAGE-1; anticancer;  
 KW melanoma; immunotherapy; immune response; variant.

XX Homo sapiens.

XX WO200023584-A1.

XX 27-APR-2000.

XX 15-OCT-1999; 99WO-EP07832.

XX 16-OCT-1998; 98EP-0119583.

XX (BOEH) BOEHRINGER INGELHEIM INT GMBH.

PA (UYHO-) UNIV HOSPITAL LEIDEN.

PI Schrier PI, Aarnoudse CA, Heider K, Klade C;

XX WPI; 2000-339685/29.

DR N-PSDB; AAD00152.

XX Tumor-associated antigen useful for cancer immunotherapy is encoded by  
 PT the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA -

XX Example 3; Page 65; 73pp; English.

XX The present sequence is the human NY-ESO-1 short variant protein, a  
 CC tumour antigen, identified by screening an esophagus carcinoma cDNA  
 CC library. This protein is derived from open reading frame (ORF)-2,  
 CC that differs from the CAMEL protein (Cytotoxic T lymphocytes (CTL)-  
 CC recognised Antigen on MELANOMA), a tumour-associated antigen, by only its  
 CC last 5 amino acids. It contains epitopes of tumour specific T-cells.  
 CC NY-ESO-1 is expressed in different tumour types, but not in healthy  
 CC tissues except in testis. The tumour-associated antigen displayed on  
 CC melanoma cells is recognised by cytotoxic T lymphocytes. This sequence  
 CC has anticancer activity. CAMEL tumour antigen and immunogenic peptides  
 CC derived from it are useful for cancer immunotherapy. They have the  
 CC potential to induce an immune response, by eliciting a CTL response.  
 CC The DNA molecule is used for the construction of recombinant or fusion  
 CC proteins.

XX Sequence 58 AA;

SQ Query Match 100.0%; Score 45; DB 21; Length 58;

Best Local Similarity 100.0%; Pred. No. 0.15;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AMLAAQERRV 10  
 DB 16 AMLAAQERRV 25

RESULT 9

AAU84819

ID AAU84819 standard; Protein; 58 AA.

XX AAU84819;

XX 08-MAY-2002 (first entry)

XX Human NYN501b consensus sequence.

XX Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;  
 KW viral infection; human immunodeficiency virus; melanoma;  
 KW bacterial infection; Salmonella; Legionella; parasitic infection;  
 KW Trypanosoma; Toxoplasma; Giardia.

XX Homo sapiens.

XX WO200190197-A1.

XX 29-NOV-2001.

XX 25-MAY-2001; 2001WO-AU00622.

XX 26-MAY-2000; 2000AU-0007761.

XX (AUSU) UNIV AUSTRALIAN NAT.

XX Thomson SA, Ramshaw IA;

XX WPI; 2002-147575/19.

XX New synthetic polypeptides having several different segments of at  
 PT least one parent polypeptide linked together differently compared to  
 PT the linkage in the parent polypeptide, for inducing immune response  
 PT against a pathogen or cancer -

XX Example 3; Fig 27; 364pp; English.

XX The invention relates to a new synthetic polypeptide (I) comprising  
 CC several different segments of at least one parent polypeptide linked  
 CC together in a different relationship relative to their linkage in the  
 CC parent polypeptide to impede, abrogate or otherwise alter at least one  
 CC function associated with the parent polypeptide and for inducing an  
 CC immune response against a pathogen or cancer. Also included are a  
 CC synthetic polynucleotide encoding and a computer system for  
 CC designing the synthetic polypeptides. The synthetic polypeptides and  
 CC polynucleotides are referred to as a savine. The synthetic polypeptide is  
 CC useful for modulating immune responses preferably directed against a  
 CC pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix,  
 CC colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone  
 CC liver, oesophagus, brain, testicle, uterus), as potentiating agents.  
 CC Compositions comprising the polypeptide may be used in the treatment or  
 CC prophylaxis against viral (such as infections caused by HIV (human  
 CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis  
 CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial  
 CC (e.g., infections caused by Neisseria, Meningococcus, Haemophilus,  
 CC Salmonella, Streptococcus, Legionella and Mycobacterium or parasitic  
 CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,  
 CC Trypanosoma, Toxoplasma and Giardia) infections. The present  
 CC sequence is a consensus sequence for a parent protein used to design a  
 CC savine of the invention.

XX Sequence 58 AA;

SQ Query Match 100.0%; Score 45; DB 23; Length 58;

Best Local Similarity 75.0%; Pred. No. 5.3e+06;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 AQEALAPL 11  
|:| | | |  
Db 1 AREALAAAL 8

Search completed: October 7, 2003, 14:23:59  
Job time : 398 secs



```
; TOPOLOGY: linear
US-08-209-204A-27
Query Match 48.0%; Score 24; DB 6; Length 9;
Best Local Similarity 75.0%; Pred. No. 5.3e+06;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 AQEALAF 11
Db 1 AKEALAA 8

RESULT 10
US-08-209-204B-27
; Sequence 27, Application US/08209204B
; GENERAL INFORMATION:
; APPLICANT: Sklar, Robert
; APPLICANT: Marchionni, Mark
; APPLICANT: Gwynne, David
; TITLE OF INVENTION: METHODS FOR TREATING MUSCLE
; TITLE OF INVENTION: DISEASES AND DISORDERS
; NUMBER OF SEQUENCES: 379
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/209,204B
; FILING DATE: 08-MAR-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/059,022
; FILING DATE: 06-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bleker-Brady, Kristina
; REGISTRATION NUMBER: 39,109
; REFERENCE/DOCKET NUMBER: 04585/028001
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-209-204B-27
Query Match 48.0%; Score 24; DB 6; Length 9;
Best Local Similarity 75.0%; Pred. No. 5.3e+06;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 AQEALAF 11
Db 1 AKEALAA 8

RESULT 11
US-08-209-204D-27
; Sequence 27, Application US/08209204D
; GENERAL INFORMATION:
; APPLICANT: Sklar, Robert
; APPLICANT: Marchionni, Mark
; APPLICANT: Gwynne, David
; TITLE OF INVENTION: METHODS FOR TREATING MUSCLE DISEASES
; TITLE OF INVENTION: DISEASES AND DISORDERS
; NUMBER OF SEQUENCES: 379
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/209,204B
; FILING DATE: 08-MAR-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/059,022
; FILING DATE: 06-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bleker-Brady, Kristina
; REGISTRATION NUMBER: 39,109
; REFERENCE/DOCKET NUMBER: 04585/028001
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-209-204D-27
Query Match 48.0%; Score 24; DB 6; Length 9;
Best Local Similarity 75.0%; Pred. No. 5.3e+06;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 AQEALAF 11
Db 1 AKEALAA 8

RESULT 12
US-08-461-097-27
; Sequence 27, Application US/08461097
; GENERAL INFORMATION:
; APPLICANT: Robert Sklar, Mark Marchionni,
; APPLICANT: David I. Gwynne
; TITLE OF INVENTION: METHODS FOR TREATING
; NUMBER OF SEQUENCES: 184
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,097
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/209,204
; FILING DATE: 08-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/059,022
; FILING DATE: 06-MAY-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 04585/028001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: 200154
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-461-097-27
Query Match 48.0%; Score 24; DB 8; Length 9;
Best Local Similarity 75.0%; Pred. No. 5.3e+06;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 AQEALAF 11
```

; PRIOR FILING DATE: 1998-10-16  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 12  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-807-512-12

Query Match 90.0%; Score 45; DB 23; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.072;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LMAQEALAF 11  
|1111111111|  
DB 1 LMAQEALAF 10

## RESULT 7

US-10-353-929-157  
; Sequence 157, Application US/10353929  
; GENERAL INFORMATION:

; APPLICANT: ITOH, Kyogo  
; TITLE OF INVENTION: Tumor antigen  
; FILE REFERENCE: GP01-1024  
; CURRENT APPLICATION NUMBER: US/10/353,929  
; CURRENT FILING DATE: 2003-01-30  
; PRIOR APPLICATION NUMBER: JP P2000-231814  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 197  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 157  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial

## ; FEATURE:

; OTHER INFORMATION: Designed peptide based on the amino acid sequence of SEQ ID NO:50

## US-10-353-929-157

Query Match 54.0%; Score 27; DB 29; Length 9;  
Best Local Similarity 75.0%; Pred. No. 5.3e-06;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LMAQEALA 9  
|:11111|  
DB 1 LLAQEAAA 8

## RESULT 8

US-08-011-396A-27  
; Sequence 27, Application US/08011396A  
; GENERAL INFORMATION:

; APPLICANT: Gwynne, David I.; Marchionni, Mark;  
; APPLICANT: McBurney, Robert N.  
; TITLE OF INVENTION: INHIBITORS OF CELL PROLIFERATION,  
; TITLE OF INVENTION: THEIR PREPARATION AND USE  
; NUMBER OF SEQUENCES: 184  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; ZIP: 02211-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

; COMPUTER: IBM

; OPERATING SYSTEM: PC-DOS

; SOFTWARE: Wordperfect

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/011,396A

; FILING DATE: 29-JAN-1993

; CLASSIFICATION: 436

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/984,085  
; FILING DATE: 01-DEC-1992  
; PRIOR APPLICATION DATA: 07/951,747  
; APPLICATION NUMBER: 07/951,747  
; FILING DATE: 25-SEP-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/927,337  
; FILING DATE: 10-AUG-1992  
; ATTORNEY/AGENT INFORMATION:

; NAME: Clark, Paul T.  
; REGISTRATION NUMBER: 30,162  
; REFERENCE/DOCKET NUMBER: 04585/017004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: 200154

; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 9

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; US-08-011-396A-27

Query Match 48.0%; Score 24; DB 4; Length 9;  
Best Local Similarity 75.0%; Pred. No. 5.3e-06;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 AQEALAF 11

|:11111|

DB 1 AKEALAL 8

## RESULT 9

US-08-209-204A-27  
; Sequence 27, Application US/08209204A  
; GENERAL INFORMATION:

; APPLICANT: Robert Sklar, Mark Marchionni,  
; APPLICANT: David I. Gwynne  
; TITLE OF INVENTION: METHODS FOR TREATING

; NUMBER OF SEQUENCES: 184

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: Massachusetts

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

; COMPUTER: IBM

; OPERATING SYSTEM: PC-DOS

; SOFTWARE: Wordperfect

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/209,204A

; FILING DATE: 08-MAR-1994

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/059,022

; FILING DATE: 06-MAY-93

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Clark, Paul T.

; REGISTRATION NUMBER: 30,162

; REFERENCE/DOCKET NUMBER: 04585/028001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 542-5070

; TELEFAX: 200154

; INFORMATION FOR SEQ ID NO: 27:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 9

; TYPE: amino acid

; STRANDEDNESS:

QY 1 MLMAQEALAEFL 11  
|||

Dd 1 MLMAQEALAEFL 11  
|||

## RESULT 2

```

PCT-US03-16736-82
; Sequence 82, Application PC/TUS0316736
; GENERAL INFORMATION:
; APPLICANT: Wang, Ronq-fu
; TITLE OF INVENTION: Mutant Fibronectin and Tumor Metastasis
; FILE REFERENCE: HO-P02484W00
; CURRENT APPLICATION NUMBER: PCT/US03/16736
; CURRENT FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/383,530
; PRIOR FILING DATE: 2002-05-28
; NUMBER OF SEQ ID NOS: 148
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 82
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
PCT-US03-16736-82

```

Query Match 100.0%; Score 50; DB 1; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0077;  
Matches 11; Conservative 0; Mismatches 0; Indels

QY            1 MLMAQEALAF L 11  
               | | | | | | | |  
Db            1 MLMAQEALAF L 11

### RESULTS

```

US-09-807-512-11
; Sequence 11, Application US/09807512
; GENERAL INFORMATION:
; APPLICANT: Schrier, Peter I.
; APPLICANT: Aarnoudse, Corlien
; APPLICANT: Heider, Karl-Heinz
; APPLICANT: Klade, Christoph
; TITLE OF INVENTION: Camel, An Alternative Translation Product of the Tumor
; TITLE OF INVENTION: Antigen-Lage 1
; FILE REFERENCE: 0652.2200000
; CURRENT APPLICATION NUMBER: US/09/807,512
; CURRENT FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: PCT/EP99/07832
; PRIOR FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: EP 98119583.7
; PRIOR FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 11
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-807-512-11

```

```
Query Match      100.0%; Score 50; DB 23; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0077;
Matches 11; Conservative 0; Mismatches 0; Indels
```

```
Qy      1 MLMAQEALAF L 11  
        | | | | |  
Db      1 MLMAQEALAF L 11
```

## RESULT A

RESOL 4  
US-10-447-161-81  
; Sequence 81, Application US/10447161

```

: GENERAL INFORMATION:
: APPLICANT: Wang, Rong-fu
: TITLE OF INVENTION: Mutant Fibronectin and Tumor Metastasis
: FILE REFERENCE: HO-P02484US1
: CURRENT APPLICATION NUMBER: US/10/447,161
: CURRENT FILING DATE: 2003-05-28
: PRIOR APPLICATION NUMBER: 60/383,530
: PRIOR FILING DATE: 2002-05-28
: NUMBER OF SEQ ID NOS: 148
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 81
: LENGTH: 11
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Synthetic Peptide
US-10-447-161-81

```

Query Match	100.0%	Score 50;	DB 30;	Length 11;
Best Local Similarity	100.0%	Pred. No. 0.0077;		
Matches \ 11;	Conservative	0;	Mismatches 0;	Indels 0;
				Gaps 0;

**QY**      1 MLMAQEALAF L 1  
           |  
           |  
**Dd**        1 MLMAQEALAF L 1

## RESULT. T 5

```

US-10-447-161-82
; Sequence 82, Application US/10447161
; GENERAL INFORMATION:
; APPLICANT: Wang, Rong-fu
; TITLE OF INVENTION: Mutant Fibronectin and Tumor Metastasis
; FILE REFERENCE: HO-P02484US1
; CURRENT APPLICATION NUMBER: US/10/447,161
; CURRENT FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/383,530
; PRIOR FILING DATE: 2002-05-28
; NUMBER OF SEQ ID NOS: 148
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 82
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-447-161-82

```

Query Match	100.0%	Score 50;	DB 30;	Length 11;
Best Local Similarity	100.0%	Pred. No. 0.0077;		
Matches 11; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy 1 MLMAQEALAF 11  
1111111111

Dy 1 MLMAOEALAF 11

## RESULT. 6

US-09-807-512-12  
US-09-807-512-12  
Sequence 12, Application US/09807512  
GENERAL INFORMATION:  
APPLICANT: Schrier, Peter I.  
APPLICANT: Aarnoudse, Corlien  
APPLICANT: Heider, Karl-Heinz  
APPLICANT: Klade, Christoph  
TITLE OF INVENTION: Camel, An Alternative Translation Product of the Tumor  
TITLE OF INVENTION: Antigen-Lage 1  
FILE REFERENCE: 0652.2200000  
CURRENT APPLICATION NUMBER: US/09/807,512  
CURRENT FILING DATE: 2001-04-16  
PRIOR APPLICATION NUMBER: PCT/EP99/07832  
PRIOR FILING DATE: 1999-10-15  
PRIOR APPLICATION NUMBER: EP 98119583.7

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:13:50 ; Search time 396 Seconds  
(without alignments)  
25.276 Million cell updates/sec

Title: US-09-807-512-11  
Perfect score: 50  
Sequence: 1 MLMAQELAPL 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues

Total number of hits satisfying chosen parameters: 624938

Minimum DB seq length: 0  
Maximum DB seq length: 11

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main.\*

- 1: /cgn2\_6/ptodata/1/paa/PCTUS\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/paa/US080\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep.\*
- 8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep.\*
- 9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep.\*
- 10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep.\*
- 11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep.\*
- 12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep.\*
- 13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep.\*
- 14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep.\*
- 15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep.\*
- 16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep.\*
- 17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep.\*
- 18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep.\*
- 19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep.\*
- 20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep.\*
- 21: /cgn2\_6/ptodata/1/paa/US097A\_COMB.pep.\*
- 22: /cgn2\_6/ptodata/1/paa/US097B\_COMB.pep.\*
- 23: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep.\*
- 24: /cgn2\_6/ptodata/1/paa/US099A\_COMB.pep.\*
- 25: /cgn2\_6/ptodata/1/paa/US099B\_COMB.pep.\*
- 26: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep.\*
- 27: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep.\*
- 28: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep.\*
- 29: /cgn2\_6/ptodata/1/paa/US103\_COMB.pep.\*
- 30: /cgn2\_6/ptodata/1/paa/US104\_COMB.pep.\*
- 31: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep.\*
- 32: /cgn2\_6/ptodata/1/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	50	100.0	11	1	PCT-US03-16736-81

Sequence 81, Appl

2	50	100.0	11	1	PCT-US03-16736-81	Sequence 82, Appl
3	50	100.0	11	23	US-09-807-512-11	Sequence 11, Appl
4	50	100.0	11	30	US-10-447-161-81	Sequence 81, Appl
5	50	100.0	11	30	US-10-447-161-81	Sequence 82, Appl
6	45	90.0	10	23	US-09-807-512-12	Sequence 12, Appl
7	27	54.0	9	29	US-10-353-929-157	Sequence 157, Appl
8	24	48.0	9	4	US-08-011-396A-27	Sequence 27, Appl
9	24	48.0	9	6	US-08-209-204A-27	Sequence 27, Appl
10	24	48.0	9	6	US-08-209-204B-27	Sequence 27, Appl
11	24	48.0	9	6	US-08-209-204B-27	Sequence 27, Appl
12	24	48.0	9	8	US-08-461-097-27	Sequence 27, Appl
13	24	48.0	9	8	US-08-461-097A-27	Sequence 27, Appl
14	24	48.0	9	8	US-08-461-097-27	Sequence 27, Appl
15	24	48.0	9	8	US-08-466-446-27	Sequence 27, Appl
16	24	48.0	9	8	US-08-467-459-27	Sequence 27, Appl
17	24	48.0	9	8	US-08-467-459A-27	Sequence 27, Appl
18	24	48.0	9	8	US-08-467-459-27	Sequence 27, Appl
19	24	48.0	9	8	US-08-467-470-27	Sequence 27, Appl
20	24	48.0	9	8	US-08-467-906-27	Sequence 27, Appl
21	24	48.0	9	8	US-08-468-731-27	Sequence 27, Appl
22	24	48.0	9	8	US-08-468-731A-27	Sequence 27, Appl
23	24	48.0	9	8	US-08-468-731-27	Sequence 27, Appl
24	24	48.0	9	8	US-08-469-526-27	Sequence 27, Appl
25	24	48.0	9	8	US-08-469-549-27	Sequence 27, Appl
26	24	48.0	9	8	US-08-469-549-27	Sequence 27, Appl
27	24	48.0	9	8	US-08-470-335-27	Sequence 27, Appl
28	24	48.0	9	8	US-08-470-335B-27	Sequence 27, Appl
29	24	48.0	9	8	US-08-470-335-27	Sequence 27, Appl
30	24	48.0	9	8	US-08-471-148-27	Sequence 27, Appl
31	24	48.0	9	8	US-08-471-833-27	Sequence 27, Appl
32	24	48.0	9	8	US-08-471-833-27	Sequence 27, Appl
33	24	48.0	9	8	US-08-471-855-27	Sequence 27, Appl
34	24	48.0	9	8	US-08-472-008-27	Sequence 27, Appl
35	24	48.0	9	8	US-08-472-065-27	Sequence 27, Appl
36	24	48.0	9	8	US-08-472-065B-27	Sequence 27, Appl
37	24	48.0	9	9	US-08-535-200-27	Sequence 27, Appl
38	24	48.0	9	9	US-08-535-200A-27	Sequence 27, Appl
39	24	48.0	9	11	US-08-734-592-27	Sequence 27, Appl
40	24	48.0	9	11	US-08-734-665-27	Sequence 27, Appl
41	24	48.0	9	11	US-08-734-665A-27	Sequence 27, Appl
42	24	48.0	9	11	US-08-734-665-27	Sequence 27, Appl
43	24	48.0	9	11	US-08-734-666-27	Sequence 27, Appl
44	24	48.0	9	11	US-08-735-010-27	Sequence 27, Appl
45	24	48.0	9	11	US-08-736-019-27	Sequence 27, Appl

ALIGNMENTS

RESULT 1  
PCT-US03-16736-81  
; Sequence 81, Application PC/TUS0316736  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Rong-fu  
; TITLE OF INVENTION: Mutant Fibronectin and Tumor Metastasis  
; FILE REFERENCE: HO-P02484W00  
; CURRENT APPLICATION NUMBER: PCT/US03/16736  
; CURRENT FILING DATE: 2003-05-28  
; PRIOR APPLICATION NUMBER: 60/383,530  
; PRIOR FILING DATE: 2002-05-28  
; NUMBER OF SEQ ID NOS: 148  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 81  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Peptide  
PCT-US03-16736-81

Query Match 100.0%; Score 50; DB 1; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0077;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



**THIS PAGE BLANK (USPTO)**

; APPLICANT: Arthur B. Raitano  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT  
; TITLE OF INVENTION: AND DETECTION OF CANCER  
; FILE REFERENCE: 129,6USU1  
; CURRENT APPLICATION NUMBER: US/09/834,765  
; CURRENT FILING DATE: 2001-09-21  
; PRIOR APPLICATION NUMBER: 60/197,647  
; PRIOR FILING DATE: 2000-04-12  
; NUMBER OF SEQ ID NOS: 770  
; SOFTWARE: FastSeq for windows Version 4.0  
; SEQ ID NO 423  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-834-765-423

Query Match 42.0%; Score 21; DB 9; Length 9;  
Best Local Similarity 62.5%; Pred. No. 5.2e+05;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 AQAALAF 11  
I:||||:|  
Db 2 AEAALSDL 9

Search completed: October 7, 2003, 14:25:42  
Job time : 66 secs

QY 1 MLMQAELAF 10  
 :| || :||  
 Db 1 LLEAQTNI AF 10

## RESULT 11

US-09-834-765-279  
 ; Sequence 279, Application US/09834765  
 ; Patent No. US20020055478A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mary Faris  
 ; APPLICANT: Pia M. Challita-Eid  
 ; APPLICANT: Arthur B. Raitano  
 ; APPLICANT: Steve Chappell Mitchell  
 ; APPLICANT: Daniel E.H. Afar  
 ; APPLICANT: Aya Jakobovits  
 ; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT  
 ; FILE REFERENCE: AND DETECTION OF CANCER  
 ; CURRENT APPLICATION NUMBER: US/09/834,765  
 ; CURRENT FILING DATE: 2001-09-21  
 ; PRIOR FILING DATE: 2000-04-12  
 ; NUMBER OF SEQ ID NOS: 770  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 279  
 ; LENGTH: 10  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-834-765-279

Query Match 44.0%; Score 22; DB 9; Length 10;  
 Best Local Similarity 50.0%; Pred. No. 3.3e+02;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MLMQAELAF 10  
 :| || :||  
 Db 1 LLEAQTNI AF 10

## RESULT 12

US-09-820-053A-140  
 ; Sequence 140, Application US/09820053A  
 ; Publication No. US20030083243A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Owen, Donald R.  
 ; TITLE OF INVENTION: SHORT BIOACTIVE PEPTIDES  
 ; FILE REFERENCE: HELX027  
 ; CURRENT APPLICATION NUMBER: US/09/820,053A  
 ; CURRENT FILING DATE: 2001-03-28  
 ; NUMBER OF SEQ ID NOS: 165  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 140  
 ; LENGTH: 11  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: SYNTHETIC SEQUENCE  
 ; NAME/KEY: MOD\_RES  
 ; LOCATION: (11)  
 ; OTHER INFORMATION: AMIDATION  
 US-09-820-053A-140

Query Match 44.0%; Score 22; DB 11; Length 11;  
 Best Local Similarity 62.5%; Pred. No. 3.6e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 MAQEALAF 10  
 :| :|||  
 Db 4 LAKLALAF 11

## RESULT 13

US-09-834-765-423

; Sequence 423, Application US/09834765  
 ; Patent No. US20020055478A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mary Faris  
 ; APPLICANT: Pia M. Challita-Eid

## US-10-109-171-140

; Sequence 140, Application US/10109171  
 ; Publication No. US20030109452A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Owen, Donald R.  
 ; TITLE OF INVENTION: SHORT BIOACTIVE PEPTIDES AND METHODS FOR THEIR USE  
 ; FILE REFERENCE: HELX028  
 ; CURRENT APPLICATION NUMBER: US/10/109,171  
 ; CURRENT FILING DATE: 2002-03-28  
 ; NUMBER OF SEQ ID NOS: 165  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 140  
 ; LENGTH: 11  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: SYNTHETIC SEQUENCE  
 ; NAME/KEY: MOD\_RES  
 ; LOCATION: (11)  
 ; OTHER INFORMATION: AMIDATION  
 US-10-109-171-140

Query Match 44.0%; Score 22; DB 15; Length 11;  
 Best Local Similarity 62.5%; Pred. No. 3.6e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 MAQEALAF 10  
 :| :|||  
 Db 4 LAKLALAF 11

## RESULT 14

US-09-834-765-53  
 ; Sequence 53, Application US/09834765  
 ; Patent No. US20020055478A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mary Faris  
 ; APPLICANT: Pia M. Challita-Eid  
 ; APPLICANT: Arthur B. Raitano  
 ; APPLICANT: Steve Chappell Mitchell  
 ; APPLICANT: Daniel E.H. Afar  
 ; APPLICANT: Aya Jakobovits  
 ; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT  
 ; FILE REFERENCE: AND DETECTION OF CANCER  
 ; CURRENT APPLICATION NUMBER: US/09/834,765  
 ; CURRENT FILING DATE: 2001-09-21  
 ; PRIOR FILING DATE: 2000-04-12  
 ; NUMBER OF SEQ ID NOS: 770  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 53  
 ; LENGTH: 9  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-834-765-53

Query Match 42.0%; Score 21; DB 9; Length 9;  
 Best Local Similarity 62.5%; Pred. No. 5.2e+05;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 AQEALAF 11  
 :| :||| :|  
 Db 2 AEEALS 9

## RESULT 15

US-09-834-765-423  
 ; Sequence 423, Application US/09834765  
 ; Patent No. US20020055478A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mary Faris  
 ; APPLICANT: Pia M. Challita-Eid

## US-09-834-765-452

Query Match 44.0%; Score 22; DB 9; Length 9;  
Best Local Similarity 62.5%; Pred. No. 5.2e+05;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 AQEALAF 11  
|| :|||  
Db 2 AQTNI AFL 9

## RESULT 7

US-09-834-765-527  
; Sequence 527, Application US/09834765  
; Patent No. US20020055478A1  
; GENERAL INFORMATION:  
; APPLICANT: Mary Faris  
; APPLICANT: Pia M. Challita-Eld  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Ava Jakobovits  
; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT  
; TITLE OF INVENTION: AND DETECTION OF CANCER  
; FILE REFERENCE: 129.6USU1  
; CURRENT APPLICATION NUMBER: US/09/834,765  
; CURRENT FILING DATE: 2001-09-21  
; PRIOR APPLICATION NUMBER: 60/197,647  
; PRIOR FILING DATE: 2000-04-12  
; NUMBER OF SEQ ID NOS: 770  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 527  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-834-765-527

Query Match 44.0%; Score 22; DB 9; Length 9;  
Best Local Similarity 62.5%; Pred. No. 5.2e+05;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 AQEALAF 11  
|| :|||  
Db 2 AQTNI AFL 9

## RESULT 8

US-09-834-765-660  
; Sequence 660, Application US/09834765  
; Patent No. US20020055478A1  
; GENERAL INFORMATION:  
; APPLICANT: Mary Faris  
; APPLICANT: Pia M. Challita-Eld  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Ava Jakobovits  
; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT  
; TITLE OF INVENTION: AND DETECTION OF CANCER  
; FILE REFERENCE: 129.6USU1  
; CURRENT APPLICATION NUMBER: US/09/834,765  
; CURRENT FILING DATE: 2001-09-21  
; PRIOR APPLICATION NUMBER: 60/197,647  
; PRIOR FILING DATE: 2000-04-12  
; NUMBER OF SEQ ID NOS: 770  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 660  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-834-765-660

Query Match 44.0%; Score 22; DB 9; Length 9;

Best Local Similarity 62.5%; Pred. No. 5.2e+05;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 AQEALAF 11  
|| :|||  
Db 2 AQTNI AFL 9

## RESULT 9

US-10-133-210-114  
; Sequence 114, Application US/10133210  
; Publication No. US20030103964A1  
; GENERAL INFORMATION:  
; APPLICANT: DeLisi, Charles  
; APPLICANT: Berzofsky, Jay  
; APPLICANT: Gulukota, Kamalakara  
; APPLICANT: Vaccaro, Dennis  
; APPLICANT: Weng, Zhiping  
; APPLICANT: Zhang, Chao  
; TITLE OF INVENTION: METHODS FOR DESIGNING MOLECULAR CONJUGATES AND  
; FILE REFERENCE: BU-035AX  
; CURRENT APPLICATION NUMBER: US/10/133,210  
; CURRENT FILING DATE: 2002-04-26  
; NUMBER OF SEQ ID NOS: 281  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 114  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-10-133-210-114

Query Match 44.0%; Score 22; DB 15; Length 9;  
Best Local Similarity 71.4%; Pred. No. 5.2e+05;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 QFALAF 11  
:| ||||  
Db 3 REDLAF 9

## RESULT 10

US-09-834-765-68  
; Sequence 68, Application US/09834765  
; Patent No. US20020055478A1  
; GENERAL INFORMATION:  
; APPLICANT: Mary Faris  
; APPLICANT: Pia M. Challita-Eld  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Ava Jakobovits  
; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT  
; TITLE OF INVENTION: AND DETECTION OF CANCER  
; FILE REFERENCE: 129.6USU1  
; CURRENT APPLICATION NUMBER: US/09/834,765  
; CURRENT FILING DATE: 2001-09-21  
; PRIOR APPLICATION NUMBER: 60/197,647  
; PRIOR FILING DATE: 2000-04-12  
; NUMBER OF SEQ ID NOS: 770  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 68  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-834-765-68

Query Match 44.0%; Score 22; DB 9; Length 10;  
Best Local Similarity 50.0%; Pred. No. 3.3e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

; APPLICANT: Arthur B. Raitano  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT  
; FILE REFERENCE: 129.6USU1  
; CURRENT APPLICATION NUMBER: US/09/834,765  
; PRIOR FILING DATE: 2001-09-21  
; PRIOR FILING DATE: 2000-04-12  
; NUMBER OF SEQ ID NOS: 770  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 176  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-834-765-176

Query Match 48.0%; Score 24; DB 9; Length 10;  
Best Local Similarity 60.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 LMAQALAEFL 11  
| | | | |  
DB 1 LEAQTNAEFL 10

RESULT 3  
US-10-353-929-28  
; Sequence 28, Application US/10353929  
; Publication No. US20030175288A1  
; GENERAL INFORMATION:  
; APPLICANT: ITOH, Kyogo  
; FILE REFERENCE: GP01-1024  
; CURRENT APPLICATION NUMBER: US/10/353,929  
; CURRENT FILING DATE: 2003-01-30  
; PRIOR APPLICATION NUMBER: JP P2000-231814  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 197  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 28  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Designed peptide having an ability to activate HLA-A2 restricted  
; OTHER INFORMATION: Cytotoxic T lymphocytes  
US-10-353-929-28

Query Match 46.0%; Score 23; DB 12; Length 9;  
Best Local Similarity 62.5%; Pred. No. 5.2e+05;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MLMAQAL 8  
: | | | |  
DB 1 LLMQLNAL 8

RESULT 4  
US-10-353-929-31  
; Sequence 31, Application US/10353929  
; Publication No. US20030175288A1  
; GENERAL INFORMATION:  
; APPLICANT: ITOH, Kyogo  
; TITLE OF INVENTION: Tumor antigen  
; FILE REFERENCE: GP01-1024  
; CURRENT APPLICATION NUMBER: US/10/353,929  
; CURRENT FILING DATE: 2003-01-30  
; PRIOR APPLICATION NUMBER: JP P2000-231814  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 197

; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 31  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Designed peptide having an ability to activate HLA-A2 restricted  
; OTHER INFORMATION: Cytotoxic T lymphocytes  
US-10-353-929-31

Query Match 46.0%; Score 23; DB 12; Length 10;  
Best Local Similarity 62.5%; Pred. No. 2.1e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MLMAQAL 8  
: | | | |  
DB 1 LLMQLNAL 8

RESULT 5  
US-10-080-797-5  
; Sequence 5, Application US/10080797  
; Publication No. US20020183253A1  
; GENERAL INFORMATION:  
; APPLICANT: Campochiaro, Peter A.  
; APPLICANT: Brazzell, Katharine H.  
; APPLICANT: Dixon, Romulus K.  
; TITLE OF INVENTION: METHOD FOR TREATING OCULAR  
; FILE REFERENCE: 4-31881A  
; CURRENT APPLICATION NUMBER: US/10/080,797  
; CURRENT FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 21  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Human  
US-10-080-797-5

Query Match 44.0%; Score 22; DB 14; Length 8;  
Best Local Similarity 100.0%; Pred. No. 5.2e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 QEALA 9  
| | | | |  
DB 4 QEALA 8

RESULT 6  
US-09-834-765-452  
; Sequence 452, Application US/09834765  
; Patent No. US20020055478A1  
; GENERAL INFORMATION:  
; APPLICANT: Mary Faris  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT  
; FILE REFERENCE: 129.6USU1  
; CURRENT APPLICATION NUMBER: US/09/834,765  
; CURRENT FILING DATE: 2001-09-21  
; PRIOR APPLICATION NUMBER: 60/197,647  
; PRIOR FILING DATE: 2000-04-12  
; NUMBER OF SEQ ID NOS: 770  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 452  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: October 7, 2003, 14:16:00 ; Search time 65 Seconds  
(without alignments)  
26.775 Million cell updates/sec

Title: US-09-807-512-11  
Perfect score: 50  
Sequence: 1 LMAQEALAF 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues

Total number of hits satisfying chosen parameters: 82306

Minimum DB seq length: 0  
Maximum DB seq length: 11

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	54.0	9	12	US-10-353-929-157
2	24	48.0	10	9	US-09-834-765-176
3	23	46.0	9	12	US-10-353-929-28
4	23	46.0	10	12	US-10-353-929-31
5	22	44.0	8	14	US-10-080-797-5
6	22	44.0	9	9	US-09-834-765-452
7	22	44.0	9	9	US-09-834-765-527
8	22	44.0	9	9	US-09-834-765-660
9	22	44.0	9	15	US-10-133-210-114
10	22	44.0	10	9	US-09-834-765-68
11	22	44.0	10	9	US-09-834-765-279
12	22	44.0	11	11	US-09-820-053A-140
13	22	44.0	11	15	US-10-109-171-140
14	21	42.0	9	9	US-09-834-765-53
15	21	42.0	9	9	US-09-834-765-423

16	21	42.0	9	9	US-09-834-765-647	Sequence 647, Appl
17	21	42.0	9	11	US-09-865-548A-98	Sequence 98, Appl
18	21	42.0	10	9	US-09-834-765-77	Sequence 77, Appl
19	21	42.0	10	9	US-09-834-765-92	Sequence 92, Appl
20	21	42.0	10	9	US-09-834-765-411	Sequence 411, Appl
21	21	42.0	10	10	US-09-995-749A-15	Sequence 15, Appl
22	21	42.0	10	11	US-09-572-404B-2781	Sequence 2781, Appl
23	21	42.0	10	12	US-09-572-270A-269	Sequence 269, Appl
24	21	42.0	10	12	US-10-353-929-30	Sequence 30, Appl
25	21	42.0	11	9	US-09-827-949-21	Sequence 21, Appl
26	20	40.0	8	9	US-09-884-681-27	Sequence 27, Appl
27	20	40.0	8	12	US-09-848-107-14	Sequence 14, Appl
28	20	40.0	8	12	US-10-293-580-27	Sequence 27, Appl
29	20	40.0	9	9	US-09-810-936-309	Sequence 309, Appl
30	20	40.0	9	10	US-09-429-755-309	Sequence 309, Appl
31	20	40.0	9	10	US-09-924-400-309	Sequence 309, Appl
32	20	40.0	9	15	US-10-212-679-309	Sequence 309, Appl
33	20	40.0	10	11	US-09-572-404B-3178	Sequence 3178, Appl
34	19	38.0	8	10	US-09-982-172-140	Sequence 140, Appl
35	19	38.0	9	8	US-08-812-393A-47	Sequence 47, Appl
36	19	38.0	9	9	US-09-834-765-558	Sequence 558, Appl
37	19	38.0	9	9	US-09-935-682-50	Sequence 50, Appl
38	19	38.0	9	9	US-09-847-185-18	Sequence 18, Appl
39	19	38.0	9	10	US-09-909-460-70	Sequence 70, Appl
40	19	38.0	9	11	US-09-277-074-10	Sequence 10, Appl
41	19	38.0	9	12	US-10-116-118-40	Sequence 40, Appl
42	19	38.0	9	12	US-10-150-797-18	Sequence 18, Appl
43	19	38.0	9	14	US-10-062-257-15	Sequence 15, Appl
44	19	38.0	9	14	US-10-106-487-18	Sequence 18, Appl
45	19	38.0	9	15	US-10-001-546-11	Sequence 11, Appl

## ALIGNMENTS

## RESULT 1

US-10-353-929-157  
; Sequence 157, Application US/10353929  
; Publication No. US20030175288A1  
; GENERAL INFORMATION:  
; APPLICANT: ITOH, Kyoto  
; TITLE OF INVENTION: Tumor antigen  
; FILE REFERENCE: GP01-1024  
; CURRENT APPLICATION NUMBER: US/10/353,929  
; CURRENT FILING DATE: 2003-01-30  
; PRIOR APPLICATION NUMBER: JP P2000-231814  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 197  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 157  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Designed peptide based on the amino acid sequence of SEQ ID NO 157  
US-10-353-929-157

Query Match 54.0%; Score 27; DB 12; Length 9;  
Best Local Similarity 75.0%; Pred. No. 5.2e+05;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LMAQEALA 9  
I:|||||  
Db 1 LLAQEAAA 8

## RESULT 2

US-09-834-765-176  
; Sequence 176, Application US/09834765  
; Patent No. US20020055478A1  
; GENERAL INFORMATION:  
; APPLICANT: Mary Faris  
; APPLICANT: Pia M. Challita-Eld

**THIS PAGE BLANK (USPTO)**

APPLICANT: Chen, Maio Su; Hiles, Ian  
TITLE OF INVENTION: Glial Mitogenic Factors, Their  
PREPARATION AND USE  
NUMBER OF SEQUENCES: 184  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
COMPUTER: IBM  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/036,555B  
FILING DATE: 24-MAR-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/965,173  
FILING DATE: 23-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/940,389  
FILING DATE: 03-SEP-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/907,138  
FILING DATE: 30-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/863,703  
FILING DATE: 03-APRIL-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: U.K. 91 07566.3  
FILING DATE: 10-APRIL-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Tsai, Christine H.  
REGISTRATION NUMBER: 34,266  
REFERENCE/DOCKET NUMBER: LUD 5250.4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
FEATURE:  
OTHER INFORMATION: Xaa in position 1 is Lysine or Arginine.  
US-08-036-555B-14

Query Match 48.0%; Score 24; DB 1; Length 10;  
Best Local Similarity 75.0%; Pred. No. 86;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 AQEALAF 11  
DB 2 AKELAAL 9

## RESULT 15

US-08-469-569-14  
Sequence 14, Application US/08469569  
Patent No. 5606032  
GENERAL INFORMATION:  
APPLICANT: Stroobant, Paul;  
APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;  
APPLICANT: Chen, Maio Su; Hiles, Ian  
TITLE OF INVENTION: Glial Mitogenic Factors, Their  
PREPARATION AND USE  
NUMBER OF SEQUENCES: 184  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
COMPUTER: IBM  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/469,569  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/036,555  
FILING DATE: 24-MAR-1993  
APPLICATION NUMBER: 07/965,173  
FILING DATE: 23-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/940,389  
FILING DATE: 03-SEP-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/907,138  
FILING DATE: 30-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/863,703  
FILING DATE: 03-APRIL-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: U.K. 91 07566.3  
FILING DATE: 10-APRIL-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Tsai, Christine H.  
REGISTRATION NUMBER: 34,266  
REFERENCE/DOCKET NUMBER: LUD 5250.4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
FEATURE:  
OTHER INFORMATION: Xaa in position 1 is Lysine or Arginine.  
US-08-469-569-14

Query Match 48.0%; Score 24; DB 1; Length 10;  
Best Local Similarity 75.0%; Pred. No. 86;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 AQEALAF 11  
DB 2 AKELAAL 9

Search completed: October 7, 2003, 14:17:15  
Job time : 29 secs



; ORGANISM: Bos taurus  
US-08-467-602-27

Query Match 48.0%; Score 24; DB 4; Length 9;  
Best Local Similarity 75.0%; Pred. No. 2.5e+05;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 AQEALAPL 11  
|:|||||  
Db 1 AKEALAL 8

## RESULT 12

PCT-US94-05083C-27  
; Sequence 27, Application PC/TUS9405083C  
; GENERAL INFORMATION:  
; APPLICANT: Robert Sklar, Mark Marchionni,  
; APPLICANT: David I. Gwyne  
; TITLE OF INVENTION: METHODS FOR ALTERING  
; TITLE OF INVENTION: MUSCLE CONDITION  
; NUMBER OF SEQUENCES: 185  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 360  
; MEDIUM TYPE: kb storage  
; COMPUTER: IBM  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/05083C  
; FILING DATE: 06-MAY-94  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/209,204  
; FILING DATE: 08-MAR-94  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/059,022  
; FILING DATE: 06-MAY-93  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Clark, Paul T.  
; REGISTRATION NUMBER: 30,162  
; REFERENCE/DOCKET NUMBER: 04585/028W01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
PCT-US94-05083C-27

Query Match 48.0%; Score 24; DB 5; Length 9;  
Best Local Similarity 75.0%; Pred. No. 2.5e+05;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 AQEALAPL 11  
|:|||||  
Db 1 AKEALAL 8

## RESULT 13

PCT-US95-06846A-27  
; Sequence 27, Application PC/TUS9506846A  
; GENERAL INFORMATION:  
; APPLICANT: Goodearl, Andrew David; Stroobant, Paul;

; APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchionni, Mark;  
; APPLICANT: Chen, Maio Su; Hiles, Ian  
; TITLE OF INVENTION: Glial Mitogenic Factors, Their  
; TITLE OF INVENTION: Preparation and Use  
; NUMBER OF SEQUENCES: 178  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Feife & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
; COMPUTER: IBM  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/06846A  
; FILING DATE: 25-MAY-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/249,322  
; FILING DATE: 26-MAY-1994  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/036,555  
; FILING DATE: 24-MAR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/965,173  
; FILING DATE: 23-OCT-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/940,389  
; FILING DATE: 03-SEP-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/907,138  
; FILING DATE: 30-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/863,703  
; FILING DATE: 03-APRIL-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: U.K. 91 07566.3  
; FILING DATE: 10-APRIL-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hanson, Norman D.  
; REGISTRATION NUMBER: 30,946  
; REFERENCE/DOCKET NUMBER: LUD 5250.5  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 688-9200  
; TELEFAX: (212) 838-3884  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
PCT-US95-06846A-27

Query Match 48.0%; Score 24; DB 5; Length 9;  
Best Local Similarity 75.0%; Pred. No. 2.5e+05;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 AQEALAPL 11  
|:|||||  
Db 1 AKEALAL 8

## RESULT 14

US-08-036-555B-14  
; Sequence 14, Application US/08036555B  
; Patent No. 5530109  
; GENERAL INFORMATION:  
; APPLICANT: Goodearl, Andrew; Stroobant, Paul;  
; APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchionni, Mark;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 1 AKEALAL 8

Qy 4 AQEALAF 11  
1:11111  
Db 1 AKEALAL 8

RESULT 9  
US-08-734-664A-27  
; Sequence 27, Application US/08734664A  
; Patent No. 6204241  
; GENERAL INFORMATION:  
; APPLICANT: Goodearl, Andrew  
; APPLICANT: Stroobant, Paul  
; APPLICANT: Minghetti, Luisa  
; APPLICANT: Waterfield, Michael  
; APPLICANT: Marchionni, Mark  
; APPLICANT: Chen, Mario  
; APPLICANT: Hiles, Ian  
; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR  
; TITLE OF INVENTION: PREPARATION AND USE  
; NUMBER OF SEQUENCES: 187  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Clark & Elbing LLP  
; STREET: 176 Federal Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IBM Compatible Pentium  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/734,664A  
; FILING DATE: 22-OCT-1996  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/249,322  
; FILING DATE: 26-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/036,555  
; FILING DATE: 24-MAR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/965,173  
; FILING DATE: 23-OCT-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/940,389  
; FILING DATE: 03-SEP-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/907,138  
; FILING DATE: 30-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/863,703  
; FILING DATE: 03-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: UK 91 07566.3  
; FILING DATE: 10-APR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bleker-Brady, Kristina  
; REGISTRATION NUMBER: 39,109  
; REFERENCE/DOCKET NUMBER: 04585/00200J  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 428-0200  
; TELEFAX: (617) 428-7045  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
US-08-734-664A-27

Query Match 48.0%; Score 24; DB 3; Length 9;  
Best Local Similarity 75.0%; Pred. No. 2.5e+05;

RESULT 10  
US-08-470-339-27  
; Sequence 27, Application US/08470339C  
; Patent No. 6232286  
; GENERAL INFORMATION:  
; APPLICANT: Goodearl, Andrew  
; APPLICANT: Stroobant, Paul  
; APPLICANT: Minghetti, Luisa  
; APPLICANT: Waterfield, Michael  
; APPLICANT: Marchionni, Mark  
; APPLICANT: Chen, Mario S.  
; APPLICANT: Hiles, Ian  
; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR  
; TITLE OF INVENTION: PREPARATION AND USE  
; FILE REFERENCE: 04585/002008  
; CURRENT APPLICATION NUMBER: US/08/470,339C  
; CURRENT FILING DATE: 1995-06-06  
; EARLIER APPLICATION NUMBER: 08/036,555  
; EARLIER FILING DATE: 1993-03-24  
; EARLIER APPLICATION NUMBER: 07/940,389  
; EARLIER FILING DATE: 1992-09-03  
; EARLIER APPLICATION NUMBER: 07/907,138  
; EARLIER FILING DATE: 1992-06-30  
; EARLIER APPLICATION NUMBER: 07/863,703  
; EARLIER FILING DATE: 1992-04-03  
; EARLIER APPLICATION NUMBER: 91 07566.3 GB  
; EARLIER FILING DATE: 1999-04-10  
; NUMBER OF SEQ ID NOS: 226  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 27  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Bos taurus  
US-08-470-339-27

Query Match 48.0%; Score 24; DB 3; Length 9;  
Best Local Similarity 75.0%; Pred. No. 2.5e+05;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 AQEALAF 11  
1:11111  
Db 1 AKEALAL 8

RESULT 11  
US-08-467-602-27  
; Sequence 27, Application US/08467602C  
; Patent No. 6444642  
; GENERAL INFORMATION:  
; APPLICANT: Sklar, Robert  
; APPLICANT: Marchionni, Mark  
; APPLICANT: Gwynne, David I.  
; TITLE OF INVENTION: METHODS FOR TREATING MUSCLE DISEASES AND  
; TITLE OF INVENTION: DISORDERS  
; FILE REFERENCE: 04585/028003  
; CURRENT APPLICATION NUMBER: US/08/467,602C  
; CURRENT FILING DATE: 1995-06-06  
; EARLIER APPLICATION NUMBER: 08/209,204  
; EARLIER FILING DATE: 1994-03-08  
; EARLIER APPLICATION NUMBER: 08/059,022  
; EARLIER FILING DATE: 1993-05-06  
; NUMBER OF SEQ ID NOS: 420  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 27  
; LENGTH: 9  
; TYPE: PRT

QY 4 AQEALAF1 11  
|:|||||  
Db 1 AKEALAL 8

## RESULT 6

US-08-469-660-27  
; Sequence 27, Application US/08469660  
; Patent No. 5876973  
; GENERAL INFORMATION:  
; APPLICANT: Gwynne, David I.; Marchionni, Mark;  
; APPLICANT: McBurrey, Robert N.  
; TITLE OF INVENTION: INHIBITORS OF CELL PROLIFERATION,  
; TITLE OF INVENTION: THEIR PREPARATION AND USE  
; NUMBER OF SEQUENCES: 184  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; ZIP: 02111-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
; COMPUTER: IBM  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/469,660  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/011,396  
; FILING DATE: 29-JAN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/984,085  
; FILING DATE: 01-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/951,747  
; FILING DATE: 25-SEP-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/927,337  
; FILING DATE: 10-AUG-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Clark, Paul T.  
; REGISTRATION NUMBER: 30,162  
; REFERENCE/DOCKET NUMBER: 04585/017004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: 200154  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
US-08-469-660-27

Query Match 48.0%; Score 24; DB 2; Length 9;  
Best Local Similarity 75.0%; Pred. No. 2.5e+05;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 AQEALAF1 11  
|:|||||  
Db 1 AKEALAL 8

## RESULT 7

US-08-470-335-27  
; Sequence 27, Application US/08470335F  
; Patent No. 6147190  
; GENERAL INFORMATION:  
; APPLICANT: GOODEARL, ANDREW

; APPLICANT: STROOBANT, PAUL  
; APPLICANT: MINGHETTI, LUISA  
; APPLICANT: WATERFIELD, MICHAEL  
; APPLICANT: MARCHIONNI, MARK  
; APPLICANT: CHEN, MARIO S.  
; APPLICANT: HILES, IAN  
; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR  
; TITLE OF INVENTION: PREPARATION AND USE  
; FILE REFERENCE: 04585/00200B  
; CURRENT APPLICATION NUMBER: US/08/470,335F  
; CURRENT FILING DATE: 1995-06-06  
; EARLIER APPLICATION NUMBER: 08/036,555  
; EARLIER FILING DATE: 1993-03-24  
; NUMBER OF SEQ ID NOS: 252  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 27  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Bos taurus  
US-08-470-335-27

Query Match 48.0%; Score 24; DB 3; Length 9;  
Best Local Similarity 75.0%; Pred. No. 2.5e+05;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 AQEALAF1 11  
|:|||||  
Db 1 AKEALAL 8

## RESULT 8

US-08-735-021-27  
; Sequence 27, Application US/08735021B  
; Patent No. 6194377  
; GENERAL INFORMATION:  
; APPLICANT: GOODEARL, ANDREW  
; APPLICANT: STROOBANT, PAUL  
; APPLICANT: MINGHETTI, LUISA  
; APPLICANT: WATERFIELD, MICHAEL  
; APPLICANT: MARCHIONNI, MARK  
; APPLICANT: CHEN, MARIO S.  
; APPLICANT: HILES, IAN  
; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR  
; TITLE OF INVENTION: PREPARATION AND USE  
; FILE REFERENCE: 04585/00200L  
; CURRENT APPLICATION NUMBER: US/08/735,021B  
; CURRENT FILING DATE: 1996-10-22  
; EARLIER APPLICATION NUMBER: 08/472,065  
; EARLIER FILING DATE: 1995-06-06  
; EARLIER APPLICATION NUMBER: 08/036,555  
; EARLIER FILING DATE: 1993-03-24  
; EARLIER APPLICATION NUMBER: 07/965,173  
; EARLIER FILING DATE: 1992-10-23  
; EARLIER APPLICATION NUMBER: 07/940,389  
; EARLIER FILING DATE: 1992-09-03  
; EARLIER APPLICATION NUMBER: 07/907,138  
; EARLIER FILING DATE: 1992-06-30  
; EARLIER APPLICATION NUMBER: 07/863,703  
; EARLIER FILING DATE: 1992-04-03  
; NUMBER OF SEQ ID NOS: 192  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 27  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Bos taurus  
US-08-735-021-27

Query Match 48.0%; Score 24; DB 3; Length 9;  
Best Local Similarity 75.0%; Pred. No. 2.5e+05;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 AQEALAF1 11  
|:|||||

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 AOEALAF11  
I:|||||  
Db 1 AKEALAL 8

## RESULT 4

US-08-469-526A-27  
; Sequence 27, Application US/08469526A  
; Patent No. 5792849

## GENERAL INFORMATION:

APPLICANT: Goodearl, Andrew  
APPLICANT: Stroobant, Paul  
APPLICANT: Minghetti, Luisa  
APPLICANT: Waterfield, Michael  
APPLICANT: Marchionni, Mark  
APPLICANT: Chen, Malo Su  
APPLICANT: Hiles, Ian

TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR

TITLE OF INVENTION: PREPARATION AND USE

NUMBER OF SEQUENCES: 187

CORRESPONDENCE ADDRESS:

ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/469,526A

FILING DATE: 06 June 1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/036,555

FILING DATE: 24-MAR-1993

APPLICATION NUMBER: 07/965,173

FILING DATE: 23-OCT-1992

APPLICATION NUMBER: 07/940,389

FILING DATE: 03-SEP-1992

APPLICATION NUMBER: 07/907,138

FILING DATE: 03-JUN-1992

APPLICATION NUMBER: 07/863,703

FILING DATE: 03-APRIL-1992

APPLICATION NUMBER: U.K. 91 07566.3

FILING DATE: 10-APR-1991

ATTORNEY/AGENT INFORMATION:

NAME: Bleker-Brady, Kristina

REGISTRATION NUMBER: 39,109

REFERENCE/DOCKET NUMBER: 04585/00200A

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-428-0200

TELEFAX: 617-428-7045

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:

LENGTH: 9

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-469-526A-27

Query Match 48.0%; Score 24; DB 1; Length 9;

Best Local Similarity 75.0%; Pred. No. 2.5e+05;

Matches 6; Conservative 1; Mismatches 1; Indels 0;

OY 4 AOEALAF11

I:|||||

Db 1 AKEALAL 8

## RESULT 5

US-08-734-591A-27  
; Sequence 27, Application US/08734591A  
; Patent No. 5854220

## GENERAL INFORMATION:

APPLICANT: Goodearl, Andrew  
APPLICANT: Stroobant, Paul  
APPLICANT: Minghetti, Luisa  
APPLICANT: Waterfield, Michael  
APPLICANT: Hiles, Ian  
APPLICANT: Marchionni, Mark  
APPLICANT: Chen, Mario

TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR

TITLE OF INVENTION: PREPARATION AND USE

NUMBER OF SEQUENCES: 187

CORRESPONDENCE ADDRESS:

ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM Compatible Pentium

OPERATING SYSTEM: Windows95

SOFTWARE: WordPerfect (Version 7.0)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/734,591A

FILING DATE: 22-OCT-1996

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/470,335

FILING DATE: 06-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/036,555

FILING DATE: 03-MAR-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/965,173

FILING DATE: 23-OCT-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/940,389

FILING DATE: 03-SEP-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/907,138

FILING DATE: 30-JUN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/863,703

FILING DATE: 03-APR-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: UK 91 07566.3

FILING DATE: 10-APR-1991

ATTORNEY/AGENT INFORMATION:

NAME: Bleker-Brady, Kristina

REGISTRATION NUMBER: 39,109

REFERENCE/DOCKET NUMBER: 04585/00200P

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 428-0200

TELEFAX: (617) 428-7045

TELEX:

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:

LENGTH: 9

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

US-08-734-591A-27

Query Match

Best Local Similarity

Matches 6; Conservative

1; Mismatches 1; Indels

0; Gaps 0;

## ; SEQUENCE CHARACTERISTICS:

; LENGTH: 9  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
US-08-036-555B-27

Query Match 48.0%; Score 24; DB 1; Length 9;  
Best Local Similarity 75.0%; Pred. No. 2.5e+05;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 AQEALAF 11  
Db 1 AKEALAL 8

## RESULT 2

US-08-469-569-27  
; Sequence 27, Application US/08469569  
; Patent No. 5606032

## ; GENERAL INFORMATION:

; APPLICANT: Goodearl, Andrew; Stroobant, Paul;  
; APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchionni, Mark;  
; APPLICANT: Chen, Maio Su; Hiles, Ian  
; TITLE OF INVENTION: Glial Mitogenic Factors, Their  
; TITLE OF INVENTION: Preparation and Use  
; NUMBER OF SEQUENCES: 184  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Felfe & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022

## ; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
; COMPUTER: IBM

## ; OPERATING SYSTEM: PC-DOS

## ; SOFTWARE: Wordperfect

## ; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/469,569

; FILING DATE: 06-JUN-1995

## ; CLASSIFICATION: 530

## ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/036,555

; FILING DATE: 24-MAR-1993

; APPLICATION NUMBER: 07/965,173

; FILING DATE: 23-OCT-1992

## ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/940,389

; FILING DATE: 03-SEP-1992

## ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/907,138

; FILING DATE: 30-JUN-1992

## ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/863,703

; FILING DATE: 03-APRIL-1992

## ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: U.K. 91 07566.3

; FILING DATE: 10-APRIL-1991

## ; ATTORNEY/AGENT INFORMATION:

; NAME: Tsai, Christine H.

; REGISTRATION NUMBER: 34,266

; REFERENCE/DOCKET NUMBER: LUD 5250.4

## ; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 688-9200

; TELEFAX: (212) 838-3884

; INFORMATION FOR SEQ ID NO: 27:

## ; SEQUENCE CHARACTERISTICS:

; LENGTH: 9

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

## US-08-469-569-27

Query Match 48.0%; Score 24; DB 1; Length 9;  
Best Local Similarity 75.0%; Pred. No. 2.5e+05;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 AQEALAF 11  
Db 1 AKEALAL 8

## RESULT 3

US-08-249-322A-27  
; Sequence 27, Application US/08249322A  
; Patent No. 5716930

## ; GENERAL INFORMATION:

; APPLICANT: Goodearl, Andrew; Stroobant, Paul;  
; APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchionni, Mark;  
; APPLICANT: Chen, Maio Su; Hiles, Ian  
; TITLE OF INVENTION: Glial Mitogenic Factors, Their  
; TITLE OF INVENTION: Preparation and Use  
; NUMBER OF SEQUENCES: 184  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Felfe & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022

## ; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
; COMPUTER: IBM

## ; OPERATING SYSTEM: PC-DOS

## ; SOFTWARE: Wordperfect

## ; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/249,322A

; FILING DATE: 26-MAY-1994

## ; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/036,555

; FILING DATE: 24-MAR-1993

; APPLICATION NUMBER: 07/965,173

; FILING DATE: 23-OCT-1992

## ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/940,389

; FILING DATE: 03-SEP-1992

## ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/907,138

; FILING DATE: 30-JUN-1992

## ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/863,703

; FILING DATE: 03-APRIL-1992

## ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: U.K. 91 07566.3

; FILING DATE: 10-APRIL-1991

## ; ATTORNEY/AGENT INFORMATION:

; NAME: Tsai, Christine H.

; REGISTRATION NUMBER: 34,266

; REFERENCE/DOCKET NUMBER: LUD 250.4

## ; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 688-9200

; TELEFAX: (212) 838-3884

; INFORMATION FOR SEQ ID NO: 27:

## ; SEQUENCE CHARACTERISTICS:

; LENGTH: 9

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

## US-08-249-322A-27

Query Match 48.0%; Score 24; DB 1; Length 9;  
Best Local Similarity 75.0%; Pred. No. 2.5e+05;

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:12:20 ; Search time 28 Seconds  
(without alignments)  
16.622 Million cell updates/sec

Title: US-09-807-512-11  
Perfect score: 50  
Sequence: 1 MLMQAEALFL 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 97567

Minimum DB seq length: 0  
Maximum DB seq length: 11

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_AA.\*  
1: /cgn2.6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2.6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2.6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2.6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2.6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2.6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	24	48.0	9	1 US-08-036-555B-27	Sequence 27, Appl
2	24	48.0	9	1 US-08-469-569-27	Sequence 27, Appl
3	24	48.0	9	1 US-08-249-322A-27	Sequence 27, Appl
4	24	48.0	9	1 US-08-469-526A-27	Sequence 27, Appl
5	24	48.0	9	2 US-08-734-591A-27	Sequence 27, Appl
6	24	48.0	9	2 US-08-469-660-27	Sequence 27, Appl
7	24	48.0	9	3 US-08-470-335-27	Sequence 27, Appl
8	24	48.0	9	3 US-08-735-021-27	Sequence 27, Appl
9	24	48.0	9	3 US-08-734-664A-27	Sequence 27, Appl
10	24	48.0	9	3 US-08-470-339-27	Sequence 27, Appl
11	24	48.0	9	4 US-08-467-602-27	Sequence 27, Appl
12	24	48.0	9	5 PCT-US94-05083C-27	Sequence 27, Appl
13	24	48.0	9	5 PCT-US95-06846A-27	Sequence 27, Appl
14	24	48.0	10	1 US-08-036-555B-14	Sequence 14, Appl
15	24	48.0	10	1 US-08-469-569-14	Sequence 14, Appl
16	24	48.0	10	1 US-08-249-322A-14	Sequence 14, Appl
17	24	48.0	10	1 US-08-469-526A-14	Sequence 14, Appl
18	24	48.0	10	2 US-08-370-909-15	Sequence 15, Appl
19	24	48.0	10	2 US-08-734-591A-14	Sequence 14, Appl
20	24	48.0	10	2 US-08-469-660-14	Sequence 14, Appl
21	24	48.0	10	3 US-08-470-335-14	Sequence 14, Appl
22	24	48.0	10	3 US-08-735-021-14	Sequence 14, Appl
23	24	48.0	10	3 US-08-734-664A-14	Sequence 14, Appl
24	24	48.0	10	3 US-08-470-339-14	Sequence 14, Appl
25	24	48.0	10	4 US-08-467-602-14	Sequence 14, Appl
26	24	48.0	10	5 PCT-US94-05083C-14	Sequence 14, Appl
27	24	48.0	10	5 PCT-US95-06846A-14	Sequence 14, Appl

28	22	44.0	9	4 US-09-434-476A-28	Sequence 28, Appl
29	22	44.0	9	4 US-09-434-476A-29	Sequence 29, Appl
30	22	44.0	9	4 US-09-434-476A-30	Sequence 30, Appl
31	22	44.0	11	1 US-08-424-957-50	Sequence 50, Appl
32	22	44.0	11	3 US-09-035-686-50	Sequence 50, Appl
33	21	42.0	9	2 US-08-408-095-4	Sequence 4, Appl
34	21	42.0	9	3 US-08-144-779C-26	Sequence 26, Appl
35	21	42.0	9	4 US-09-368-449B-26	Sequence 26, Appl
36	21	42.0	9	4 US-09-434-476A-31	Sequence 31, Appl
37	21	42.0	10	4 US-09-604-957-9	Sequence 9, Appl
38	21	42.0	11	1 US-08-528-199-4	Sequence 4, Appl
39	21	42.0	11	1 US-08-542-363-21	Sequence 21, Appl
40	21	42.0	11	3 US-09-100-089-21	Sequence 21, Appl
41	21	42.0	11	4 US-09-670-827-21	Sequence 21, Appl
42	20	40.0	6	1 US-08-260-202A-34	Sequence 34, Appl
43	20	40.0	8	2 US-08-679-865-27	Sequence 27, Appl
44	20	40.0	8	2 US-08-680-876-27	Sequence 27, Appl
45	20	40.0	8	3 US-09-263-975-27	Sequence 27, Appl

## ALIGNMENTS

## RESULT 1

US-08-036-555B-27  
; Sequence 27, Application US/08036555B  
; Patent No. 5530109  
; GENERAL INFORMATION:  
; APPLICANT: Goodearl, Andrew; Stroobant, Paul;  
; APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;  
; APPLICANT: Chen, Maio Su; Hiles, Ian  
; TITLE OF INVENTION: Glial Mitogenic Factors, Their  
; TITLE OF INVENTION: Preparation and Use  
; NUMBER OF SEQUENCES: 184  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Felfe & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
; COMPUTER: IBM  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/036.555B  
; FILING DATE: 24-MAR-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/965,173  
; FILING DATE: 23-OCT-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/940,389  
; FILING DATE: 03-SEP-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/907,138  
; FILING DATE: 30-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/863,703  
; FILING DATE: 03-APRIL-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: U.K. 91 07566.3  
; FILING DATE: 10-APRIL-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Tsai, Christine H.  
; REGISTRATION NUMBER: 34,266  
; REFERENCE/DOCKET NUMBER: LUD 5250.4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 688-9200  
; TELEFAX: (212) 838-3884  
; INFORMATION FOR SEQ ID NO: 27:

**THIS PAGE BLANK (USPTO)**

GN GLK.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PA01;  
RX MEDLINE=9642734; PubMed=8830708;  
RA Sage A.E., Proctor W.D., Phipps P.V.Jr.;  
RT "A two-component response regulator, glrR, is required for glucose  
RL transport activity in Pseudomonas aeruginosa PA01.";  
RL J. Bacteriol. 178:6064-6066(1996).  
DR EMBL; U50932; AAC44474.1; -.  
KW Kinase.  
FT NON\_TER 1  
SQ SEQUENCE 9 AA; 944 MW; C3071DDAA72DC6C6 CRC64;  
  
Query Match 30.0%; Score 15; DB 2; Length 9;  
Best Local Similarity 75.0%; Pred. NO. 8.3e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 5 QEAL 8  
Db 2 QOAL 5  
  
Search completed: October 7, 2003, 14:15:55  
Job time : 97 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:04:19 ; Search time 94 Seconds  
(without alignments)  
30.198 Million cell updates/sec

Title: US-09-807-512-11

Perfect score: 50

Sequence: 1 MLMQAQALFL 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 1748

Minimum DB seq length: 0  
Maximum DB seq length: 11

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_23:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phage:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_rvirus:\*

16: sp\_bacteriap:\*

17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	40.0	11	4 Q9UE69	Q9ue69 homo sapien
2	20	40.0	11	7 Q9TQB3	Q9tqb3 homo sapien
3	18.5	37.0	9	2 Q46179	Q46179 clostridium
4	18	36.0	9	13 Q8AYL5	Q8ayl5 carassius a
5	18	36.0	10	2 Q52762	Q52762 rhizobium s
6	18	36.0	10	12 Q66190	Q66190 avian infec
7	17	34.0	8	2 P83152	P83152 anabaena sp
8	17	34.0	10	2 Q8RIT1	Q8rit1 anaplasma p
9	17	34.0	11	3 Q60005	Q60005 aspergillus
10	17	34.0	11	3 Q60007	Q60007 emericeila
11	17	34.0	11	3 Q60192	Q60192 aspergillus
12	17	34.0	11	3 Q60006	Q60006 aspergillus
13	16	32.0	11	2 Q44237	Q44237 anabaena sp
14	16	32.0	11	12 Q84247	Q84247 poliovirinu
15	15	30.0	9	2 P72149	P72149 pseudomonas
16	15	30.0	9	2 P83157	P83157 anabaena sp

17 15 30.0 9 4 Q9NYH5 Q9nyh5 homo sapien

18 15 30.0 9 5 Q9TVF1 Q9tvf1 trypanosoma

19 15 30.0 9 13 Q8AUM7 Q8aum7 carassius a

20 15 30.0 10 2 Q48469 Q48469 klebsiella

21 15 30.0 10 10 Q8LRW7 Q8lrt7 chlamydomon

22 15 30.0 10 11 Q70580 Q70580 mus musculu

23 15 30.0 11 2 Q9EUZ3 Q9euz3 escherichia

24 15 30.0 11 10 Q06626 Q06626 solanum tub

25 14 28.0 9 6 Q8MJT7 Q8mit7 eulemur ful

26 14 28.0 9 6 Q8MJT8 Q8mit8 eulemur ful

27 14 28.0 9 10 Q9AXH8 Q9axh8 mesembryant

28 14 28.0 9 12 Q69100 Q69100 herpes simp

29 14 28.0 10 1 Q9UWM5 Q9uwm5 sulfolobus

30 14 28.0 10 2 Q56097 Q56097 salmonella

31 14 28.0 10 2 Q9R7J8 Q9r7j8 helicobacte

32 14 28.0 10 6 Q9TQV4 Q9tcv4 equus cabal

33 14 28.0 10 11 Q8CJF0 Q8cje0 rattus norv

34 14 28.0 10 11 Q8BHN2 Q8bhn2 mus musculu

35 14 28.0 10 12 Q9Q100 Q9qlu0 saimirine

36 14 28.0 10 12 P90373 P90373 pseudorabie

37 14 28.0 10 13 Q73588 Q73588 gallus gall

38 14 28.0 11 3 Q96V15 Q96v15 cryptococcu

39 14 28.0 11 6 Q9TQS9 Q9tqs9 equus cabal

40 14 28.0 11 7 Q77912 Q77912 oreochromis

41 14 28.0 11 7 Q77883 Q77883 oreochromis

42 14 28.0 11 13 Q9PS71 Q9ps71 agkistrodon

43 13 26.0 8 15 Q98YK9 Q98yk9 human immun

44 13 26.0 9 4 Q9BX14 Q9bx14 homo sapien

45 13 26.0 9 10 P82440 P82440 nicotiana t

#### ALIGNMENTS

##### RESULT 1

Q9UE69 PRELIMINARY; PRT; 11 AA.

AC Q9UE69; DT 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)

DE 5HT3 serotonin receptor (Fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Bedford F.K., Taylor S., Julius D., Ingraham H.A.;

RT "Expression of the 5HT3 serotonin receptor gene in neuronal cells is regulated via novel NF-1 mediated complexes.";

RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL; U73443; AAB71736.1; -.

KW Receptor.

FT NON\_TER 11

SQ SEQUENCE 11 AA; 1286 MW; 28F50414C6C6C2D3 CRC64;

Query Match 40.0%; Score 20; DB 4; Length 11;  
Best Local Similarity 50.0%; Pred. No. 2.2e+03;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MLMQAQAL 8  
DB 2 LLWVQVAL 9

##### RESULT 2

Q9TQB3 PRELIMINARY; PRT; 11 AA.

ID Q9TQB3

DT 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE MHC class I related protein 1 (Fragment).  
GN MR1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98451457; PubMed=9780177;  
RA Riegert P., Wanner V., Bahrer S.;  
RT "Genomics, isoforms, expression, and phylogeny of the MHC class I-  
related MR1 gene."  
RL J. Immunol. 161:4066-4077(1998).  
DR EMBL; AF039526; AA002172.1; -.  
FT NON\_TER 11  
SQ SEQUENCE 11 AA; 1235 MW; 5E71A31E29CDD697 CRC64;  
  
Query Match 40.0%; Score 20; DB 7; Length 11;  
Best Local Similarity 66.7%; Pred. No. 2.2e+03;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 6 EALAF 11  
DB 3 ELMAFL 8  
  
RESULT 3  
Q46179  
ID Q46179 PRELIMINARY; PRT; 9 AA.  
AC Q46179;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Chloramphenicol acetyltransferase.  
GN CATQ.  
OS Clostridium perfringens.  
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
OC Clostridium.  
OX NCBI\_TaxID=1502;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=CW 531;  
RX MEDLINE=91247774; PubMed=2039197;  
RA Bannam T.L., Rood J.I.;  
RT "The relationship between the Clostridium perfringens catQ gene  
product and chloramphenicol acetyltransferases from other bacteria."  
RL Antimicrob. Agents Chemother. 35:471-476(1991).  
DR EMBL; M55620; AAA23214.1; -.  
KW Transferase.  
SQ SEQUENCE 9 AA; 1041 MW; AFF4D72322CDD696 CRC64;  
  
Query Match 37.0%; Score 18.5; DB 2; Length 9;  
Best Local Similarity 60.0%; Pred. No. 8.3e+05;  
Matches 6; Conservative 2; Mismatches 1; Indels 1; Gaps 1;  
  
QY 1 MLMQAEALAF 10  
DB 1 MMMAVK-LAF 9  
  
RESULT 4  
Q8AVL5  
ID Q8AVL5 PRELIMINARY; PRT; 9 AA.  
AC Q8AVL5;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Cytochrome P450 aromatase (Fragment).  
GN CYP19A.  
OS Carassius auratus (Goldfish).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Carassius.

OX NCBI\_TaxID=7957;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Tchoudakova A.V., Kishida M., Wood E., Callard G.V.;  
RT "Promoter characteristics of two Cyp19 genes differentially expressed  
in the brain and ovary of teleost fish."  
RL J. Steroid Biochem. Mol. Biol. 0:0-0(2001).  
DR EMBL; AF324897; AAN32618.1; -.  
FT NON\_TER 9  
SQ SEQUENCE 9 AA; 1060 MW; C49E76D7272B040D CRC64;  
  
Query Match 36.0%; Score 18; DB 13; Length 9;  
Best Local Similarity 66.7%; Pred. No. 8.3e+05;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 3 MAQEAL 8  
DB 1 MARELL 6  
  
RESULT 5  
Q52762  
ID Q52762 PRELIMINARY; PRT; 10 AA.  
AC Q52762;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Dinitrogenase beta-subunit (Fragment).  
OS Rhizobium sp. cowpea (strain Irc78).  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Rhizobiaceae; Rhizobium.  
OX NCBI\_TaxID=400;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=Irc78;  
RA Yun A.C., Szalay A.A.;  
RT "Structural genes of dinitrogenase and dinitrogenase reductase are  
transcribed from two separate promoters in the broad host range cowpea  
Rhizobium strain Irc78."  
RL Proc. Natl. Acad. Sci. U.S.A. 81:7358-7362(1984).  
DR EMBL; M10203; AAA26309.1; -.  
FT NON\_TER 10  
SQ SEQUENCE 10 AA; 1100 MW; CFFD08C1EB1DC5A6 CRC64;  
  
Query Match 36.0%; Score 18; DB 2; Length 10;  
Best Local Similarity 80.0%; Pred. No. 5.4e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 3 MAQEA 7  
DB 1 MAQSA 5  
  
RESULT 6  
Q66190  
ID Q66190 PRELIMINARY; PRT; 10 AA.  
AC Q66190;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Hypothetical 1.1 kDa protein (Fragment).  
OS Avian infectious bronchitis virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
OC Coronaviridae; Coronavirus.  
OX NCBI\_TaxID=11120;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=41;  
RX MEDLINE=87021475; PubMed=2429473;  
RA Niesters H.G., Lenstra J.A., Spaan W.J., Zijderveld A.J.,  
RA Bleumink-Pluym N.M., Hong F., van Scharrenburg G.J., Horzinek M.C.,  
RA van der Zeijst B.A.;  
RT "The peplomer protein sequence of the M41 strain of coronavirus IBV

RT and its comparison with Beaudette strains.";

RL Virus Res. 5:253-263(1986).

DR EMBL; M21883; AAA66576.1; -.

KW Hypothetical protein.

FT NON\_TER 10

SQ SEQUENCE 10 AA; 1136 MW; CF2510D5A1B775A6 CRC64;

Query Match 36.0%; Score 18; DB 12; Length 10;

Best Local Similarity 44.4%; Pred. No. 5.4e+03;

Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 MAQALAFLL 11

DB 1 MIQSPTSFL 9

RESULT 7

P83152

ID P83152 PRELIMINARY; PRT; 8 AA.

AC P83152;

DT 01-MAR-2002 (TREMBLrel. 20, Created)

DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE Allophycocyanin beta chain (Fragment).

OS Anabaena sp. (strain L31).

OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.

OX NCBI\_TaxID=29412;

RN [1]

RP SEQUENCE.

RA Apte S.K., Uhlemann E., Schmid R., Altendorf K.;

RL Submitted (OCT-2001) to the SWISS-PROT data bank.

CC -!- FUNCTION: LIGHT-HARVESTING PHOTOSYNTHETIC BILE PIGMENT-PROTEIN

FROM THE PHYCOBILIPROTEIN COMPLEX. ALLOPHYCOCYANIN HAS A MAXIMUM

CC ABSORPTION AT APPROXIMATELY 650 TO 653 NANOMETERS.

CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.

CC -!- SUBCELLULAR LOCATION: THYLAKOID MEMBRANE.

CC -!- PTM: CONTAINS ONE COVALENTLY LINKED BILIN CHROMOPHORE.

KW Phycobilisome; Electron transport; Photosynthesis; Bile pigment;

KW Thylakoid; Membrane.

FT NON\_TER 8

SQ SEQUENCE 8 AA; 788 MW; 87CDC1A05DDAB56DD CRC64;

Query Match 34.0%; Score 17; DB 2; Length 8;

Best Local Similarity 60.0%; Pred. No. 8.3e+05;

Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 AQEAL 8

DB 1 AQDAI 5

RESULT 8

Q8RIT1

ID Q8RIT1 PRELIMINARY; PRT; 10 AA.

AC Q8RIT1;

DT 01-JUN-2002 (TREMBLrel. 21, Created)

DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Heat shock protein groES (Fragment).

GN GROES.

OS Anaplasma phagocytophilum (Ehrlichia phagocytophila).

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;

OC Anaplasmataceae; Anaplasma.

OX NCBI\_TaxID=948;

RN [1]

RP SEQUENCE FROM N.A.

RA von Loewenich F.D., Bodan C.;

RT "A case of equine granulocytic ehrlichiosis provides further evidence

for the presence of the HGE agent in Germany."

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF482760; AAL88676.1; -.

FT NON\_TER 1

SQ SEQUENCE 10 AA; 1103 MW; 9792B43DD0505AB5 CRC64;

Query Match 34.0%; Score 17; DB 2; Length 10;

Best Local Similarity 50.0%; Pred. No. 8.7e+03;

Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LMAQEA 7

DB 5 IIAKEA 10

RESULT 9

O60005

ID O60005 PRELIMINARY; PRT; 11 AA.

AC O60005;

DT 01-AUG-1998 (TREMBLrel. 07, Created)

DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE TrpC (Fragment).

GN TRPC.

OS Aspergillus versicolor.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;

OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.

OX NCBI\_TaxID=46472;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NRRL 226;

RA Geiser D.M., Taylor J.W., Smith G.W., Ritchie K.B.;

RT "Aspergillus sydowii causing sea fan mortality."

RL Nature 0:0-0(1998).

DR EMBL; AF058967; AAC15743.1; -.

FT NON\_TER 11

SQ SEQUENCE 11 AA; 1142 MW; 8C71EBD3B2C72DC5 CRC64;

Query Match 34.0%; Score 17; DB 3; Length 11;

Best Local Similarity 66.7%; Pred. No. 9.6e+03;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 MAQEAL 8

DB 1 MADSAAL 6

RESULT 10

O60007

ID O60007 PRELIMINARY; PRT; 11 AA.

AC O60007;

DT 01-AUG-1998 (TREMBLrel. 07, Created)

DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE TrpC (Fragment).

GN TRPC.

OS Emericella violacea.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;

OC Eurotiales; Trichocomaceae; Emericella.

OX NCBI\_TaxID=41738;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 16813;

RA Geiser D.M., Taylor J.W., Smith G.W., Ritchie K.B.;

RT "Aspergillus sydowii causing sea fan mortality."

RL Nature 0:0-0(1998).

DR EMBL; AF058975; AAC15751.1; -.

FT NON\_TER 11

SQ SEQUENCE 11 AA; 1142 MW; 8C71EBD3B2C72DC5 CRC64;

Query Match 34.0%; Score 17; DB 3; Length 11;

Best Local Similarity 66.7%; Pred. No. 9.6e+03;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 MAQEAL 8

DB 1 MADSAAL 6

## RESULT 11

O60192 ID O60192 PRELIMINARY; PRT; 11 AA.  
 AC O60192;  
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)  
 DE TRPC (Fragment).  
 GN TRPC.  
 OS Aspergillus sydowii.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.  
 OX NCBI\_TaxID=75750;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NRRL 249, H640, SA-25, SS-7, FK-11, and NRRL 244;  
 RA Geiser D.M., Taylor J.W., Smith G.W., Ritchie K.B.;  
 RL Nature 0:0-0(1998).  
 DR EMBL; AF058974; AAC15750.1; -;  
 DR EMBL; AF058968; AAC15744.1; -;  
 DR EMBL; AF058969; AAC15745.1; -;  
 DR EMBL; AF058970; AAC15746.1; -;  
 DR EMBL; AF058971; AAC15747.1; -;  
 DR EMBL; AF058973; AAC15749.1; -;  
 FT NON\_TER 11  
 SQ SEQUENCE 11 AA; 1142 MW; 8C71EBD3B2C72DC5 CRC64;

Query Match 34.0%; Score 17; DB 3; Length 11;  
 Best Local Similarity 66.7%; Pred. No. 9.6e+03;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 MAQEAL 8  
 Db 1 MADSAL 6

## RESULT 12

O60006 ID O60006 PRELIMINARY; PRT; 11 AA.  
 AC O60006;  
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE TRPC (Fragment).  
 GN TRPC.  
 OS Aspergillus sydowii.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.  
 OX NCBI\_TaxID=75750;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NRRL 242;  
 RA Geiser D.M., Taylor J.W., Smith G.W., Ritchie K.B.;  
 RT "Aspergillus sydowii causing sea fan mortality.";  
 RL Nature 0:0-0(1998).  
 DR EMBL; AF058972; AAC15748.1; -;  
 FT NON\_TER 11  
 SQ SEQUENCE 11 AA; 1110 MW; 8C71F0C3F2C72DC5 CRC64;

Query Match 34.0%; Score 17; DB 3; Length 11;  
 Best Local Similarity 66.7%; Pred. No. 9.6e+03;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 MAQEAL 8  
 Db 1 MADSAL 6

## RESULT 13

O44237 ID O44237 PRELIMINARY; PRT; 11 AA.  
 AC O44237;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Glutamine synthetase (Fragment).  
 GN GLNA.  
 OS Anabaena sp. (strain PCC 7120).  
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
 OX NCBI\_TaxID=103690;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PCC 7120;  
 RA Warner L.E., Ligon P.J., Stahel A.W., Curtis S.E.;  
 RT "The apcF gene of Anabaena sp. strain PCC 7120 is regulated by  
 RT nitrogen and the apcF and glnA promoters overlap.";  
 RL Submitted (MAY-1995) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PCC 7120;  
 RA Scappino L.A.;  
 RL Submitted (FEB-1995) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; U21853; AAA65652.1; -;  
 FT NON\_TER 11  
 SQ SEQUENCE 11 AA; 1316 MW; 2000580E32CB06C7 CRC64;

Query Match 32.0%; Score 16; DB 2; Length 11;  
 Best Local Similarity 50.0%; Pred. No. 1.6e+04;  
 Matches 4; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MLMAQEAL 8  
 Db 1 MTTPOEVL 8

## RESULT 14

Q84247 ID Q84247 PRELIMINARY; PRT; 11 AA.  
 AC Q84247;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Capsid protein VP2 (Fragment).  
 OS Polyomavirus BK (BKV).  
 OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.  
 OX NCBI\_TaxID=10629;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90324932; PubMed=2165132;  
 RA Moens U., Sundsfjord A., Flegstad T., Traavik T.;  
 RT "BK virus early RNA transcripts in stably transformed cells: enhanced  
 RT levels induced by dibutyryl cAMP, forskolin and 12-O-  
 RT tetradecanoylphorbol-13-acetate treatment.";  
 RL J. Gen. Virol. 71:1461-1471(1990).  
 DR EMBL; D00678; BAA00585.1; -;  
 FT NON\_TER 11  
 SQ SEQUENCE 11 AA; 1044 MW; C2786C4E272DD72D CRC64;

Query Match 32.0%; Score 16; DB 12; Length 11;  
 Best Local Similarity 80.0%; Pred. No. 1.6e+04;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 7 ALAFL 11  
 Db 4 ALALL 8

## RESULT 15

P72149 ID P72149 PRELIMINARY; PRT; 9 AA.  
 AC P72149;  
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Putative glucokinase (Fragment).

**THIS PAGE BLANK (USPTO)**



RL J. Bacteriol. 174:2606-2611(1992).  
 CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA AND IS  
 CC LOCATED AT THE A SITE OF THE PEPTIDYLTRANSFERASE CENTER  
 CC (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE L16P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M74771; AAA21914.1; -.  
 DR PIR: F41839; F41839.  
 DR InterPro: IPR000114; Ribosomal\_L16.  
 DR PROSITE: PS00586; RIBOSOMAL\_L16\_1; PARTIAL.  
 DR PROSITE: PS00701; RIBOSOMAL\_L16\_2; PARTIAL.  
 KW Ribosomal protein; rRNA-binding.  
 FT NON\_TER 10 10  
 SQ SEQUENCE 10 AA; 1324 MW; B3386A21B4032766 CRC64;

Query Match 28.0%; Score 14; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 5.3e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLM 3  
 |||  
 Db 1 MLM 3

## RESULT 10

ALL3\_CARMA  
 ID ALL3\_CARMA STANDARD; PRT; 7 AA.  
 AC P81806;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Carcinustatin 3.  
 OS Carcinus maenas (Common shore crab) (Green crab).  
 CC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 CC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
 CC Eubrachyura; Portunoidae; Portunidae; Carcinus.  
 OX NCBI\_TaxID=6759;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;  
 RX MEDLINE=98121193; PubMed=9461295;  
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,  
 RA Thorpe A.;  
 RT "Isolation and identification of multiple neuropeptides of the  
 RT allatostatin superfamily in the shore crab Carcinus maenas.";  
 RL Eur. J. Biochem. 250:727-734(1997).  
 CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.  
 CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
 KW Neuropeptide; Multigene family.  
 SQ SEQUENCE 7 AA; 796 MW; 672879CDCB476B70 CRC64;

Query Match 26.0%; Score 13; DB 1; Length 7;  
 Best Local Similarity 60.0%; Pred. No. 1.3e+05;  
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 EALAF 10  
 |||  
 Db 1 EYAF 5

## RESULT 11

COXG\_RAT  
 ID COXG\_RAT STANDARD; PRT; 8 AA.  
 AC P80430;  
 DT 01-NOV-1995 (Rel. 32, Created)

QY 4 AQEALAF 10  
 | : ||

DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 28-FEB-2000 (Rel. 41, Last annotation update)  
 DE Cytochrome c oxidase polypeptide VIB (EC 1.9.3.1) (AED) (Fragment).  
 GN COX6B.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=Wistar; TISSUE=Liver;  
 RX MEDLINE=95324529; PubMed=7601105;  
 RA Schaeffer H., Noack H., Halangk W., Brandt U., von Jagow G.;  
 RT "Cytochrome-c oxidase in developing rat heart. Enzymic properties and  
 RT amino-terminal sequences suggest identity of the fetal heart and the  
 RT adult liver isoform".  
 RL Eur. J. Biochem. 230:235-241(1995).  
 CC -1- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE  
 CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN  
 CC MITOCHONDRIAL ELECTRON TRANSPORT. THIS PROTEIN MAY BE ONE OF THE  
 CC HEME-BINDING SUBUNITS OF THE OXIDASE.  
 CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome  
 CC c + 2 H(2)O.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VIB FAMILY.  
 DR PIR: S65381; S65381.  
 KW Oxidoreductase; Mitochondrion.  
 FT NON\_TER 1 1  
 FT NON\_TER 8 8  
 SQ SEQUENCE 8 AA; 1039 MW; 8101E9CAA73AE456 CRC64;

Query Match 26.0%; Score 13; DB 1; Length 8;  
 Best Local Similarity 50.0%; Pred. No. 1.3e+05;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 QEALAF 10  
 | | | |  
 Db 1 QNXLDF 6

## RESULT 12

ALL10\_CARMA  
 ID ALL10\_CARMA STANDARD; PRT; 9 AA.  
 AC P81813;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Carcinustatin 10.  
 OS Carcinus maenas (Common shore crab) (Green crab).  
 CC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 CC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
 CC Eubrachyura; Portunoidae; Portunidae; Carcinus.  
 OX NCBI\_TaxID=6759;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;  
 RX MEDLINE=98121193; PubMed=9461295;  
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,  
 RA Thorpe A.;  
 RT "Isolation and identification of multiple neuropeptides of the  
 RT allatostatin superfamily in the shore crab Carcinus maenas.";  
 RL Eur. J. Biochem. 250:727-734(1997).  
 CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.  
 CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
 KW Neuropeptide; Multigene family.  
 FT MOD\_RES 9 9  
 SQ SEQUENCE 9 AA; 963 MW; 372D79DCB4776C7 CRC64;

Query Match 26.0%; Score 13; DB 1; Length 9;  
 Best Local Similarity 42.9%; Pred. No. 1.3e+05;  
 Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 AQEALAF 10  
 | : ||

```

RESULT 6
FIBB_THEGE
ID FIBB_THEGE STANDARD; PRT; 9 AA.
AC P19342;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).
GN FGB.
OS Theropithecus gelada (Gelada baboon).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Theropithecus.
OX NCBI_TaxID=9565;
RN [1]
RP SEQUENCE.
RX MEDLINE=84161822; PubMed=6423621;
RA Nakamura S., Takenaka O., Takahashi K.;
RT "Fibrinopeptides A and B of baboons (Papio anubis, Papio hamadryas,
RT and Theropithecus gelada): their amino acid sequences and
RT evolutionary rates and a molecular phylogeny for the baboons.";
RL J. Biochem. 94:1973-1978(1983).
CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC AGGREGATION.
CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
CC PIR: F28854; F28854.
DR InterPro: IPR002191; Fibrinogen_C.
DR PROSITE: PS00514; FIBRINAG_C_DOMAIN; PARTIAL.
KW Blood coagulation; Plasma.
FT PEPTIDE 1 9 FIBRINOPEPTIDE B.
FT NON_TER 1 9
SQ SEQUENCE 9 AA; 977 MW; DDPE7879C7287B06 CRC64;

Query Match 28.0%; Score 14; DB 1; Length 9;
Best Local Similarity 75.0%; Pred. No. 1.3e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 QEAL 8
Db 2 QEGL 5

RESULT 7
MOSF_CLYJA
ID MOSF_CLYJA STANDARD; PRT; 9 AA.
AC P19853;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE [Phe-6]-mosact.
OS Clypeaster japonicus (Sand dollar).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Gnathostomata; Clypeasteroidea;
OC Clypeasteridae; Clypeaster.
OX NCBI_TaxID=7644;
RN [1]
RP SEQUENCE.
RC TISSUE=Egg jelly;
RA Suzuki N., Kurita M., Yoshino K.I., Kajiuira H., Nomura K.,
RT "Purification and structure of mosact and its derivatives from the
RT egg jelly of the sea urchin Clypeaster japonicus.";
RL Zool. Sci. 4:649-656(1987).
CC -1- FUNCTION: Stimulates sperm respiration and motility.
DR PIR: JN0027; JN0027.
SQ SEQUENCE 9 AA; 924 MW; 93245729CDC5BAB5 CRC64;

```

```

Query Match 28.0%; Score 14; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 AFL 11
Db 5 AFL 7

RESULT 8
Q2OG_COMTE
ID Q2OG_COMTE STANDARD; PRT; 10 AA.
AC P80466;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Quinolone 2-oxidoeductase, gamma chain (EC 1.3.99.17) (Fragment).
OS Comamonas testosteroni (Pseudomonas testosteroni).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Comamonadaceae; Comamonas.
OX NCBI_TaxID=285;
RN [1]
RP SEQUENCE.
RC STRAIN=63;
RX MEDLINE=96035889; PubMed=7556204;
RT "Quinolone 2-oxidoeductase and 2-oxo-1,2-dihydroquinoline 5,6-
RT dioxygenase from Comamonas testosteroni 63. The first two enzymes in
RT quinoline and 3-methylquinoline degradation.";
RL Eur. J. Biochem. 232:536-544(1995).
CC -1- FUNCTION: CONVERTS (3-METHYL-)-QUINOLINE TO (3-METHYL-)-2-OXO-
CC 1,2-DIHYDROQUINOLINE.
CC -1- CATALYTIC ACTIVITY: Quinolone + acceptor + H(2)O = isoquinolin-
CC 1(2H)-one + reduced acceptor.
CC -1- COFACTOR: FAD, MOLYBDENUM AND IRON-SULFUR.
CC -1- PATHWAY: Degradation of quinoline and (3-methyl-)-quinoline; first
CC step.
CC -1- SUBUNIT: HETEROHEXAMER OF TWO ALPHA CHAINS, TWO BETA CHAINS, AND
CC TWO GAMMA CHAINS (PROBABLE).
KW Oxidoreductase; Flavoprotein; FAD; Molybdenum.
FT NON_TER 10 10
SQ SEQUENCE 10 AA; 1153 MW; C848CE64433B1DC6 CRC64;

Query Match 28.0%; Score 14; DB 1; Length 10;
Best Local Similarity 33.3%; Pred. No. 5.3e+03;
Matches 2; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLMAQE 6
Db 1 MIQAEK 6

RESULT 9
RL16_ACHLA
ID RL16_ACHLA STANDARD; PRT; 10 AA.
AC P29221;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE 50S ribosomal protein L16 (Fragment).
GN RPLP.
OS Acholeplasma laidlawii.
OC Bacteria; Firmicutes; Mollicutes; Acholeplasmatales;
OC Acholeplasmataceae; Acholeplasma.
OX NCBI_TaxID=2148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92210505; PubMed=1556079;
RA Lim P.O., Sears B.B.;
RT "Evolutionary relationships of a plant-pathogenic mycoplasma-like
RT organism and Acholeplasma laidlawii deduced from two ribosomal protein
RT gene sequences.";

```



```

RT "The small subunit of the mammalian mitochondrial ribosome:
RT identification of the full complement of ribosomal proteins present.";
RL J. Biol. Chem. 276:19363-19374(2001).
CC -1- SUBUNIT: Component of the mitochondrial ribosome small subunit
CC (28S) which comprises a 12S rRNA and about 30 distinct proteins.
CC -1- SUBCELLULAR LOCATION: Mitochondrial.
KW Ribosomal protein; Mitochondrion.
FT NON_TER 1 1
FT NON_TER 9 9
SQ SEQUENCE 9 AA; 1032 MW; D341D73776DB05B9 CRC64;

Query Match 32.0%; Score 16; DB 1; Length 9;
Best Local Similarity 37.5%; Pred. No. 1.3e+05;
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 LMAQEALA 9
DB 1 LFEQEPLA 8
I : : : I I

RESULT 3
CLP_THICU STANDARD; PRT; 8 AA.
AC P80488;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Chemolithotroph-specific protein (Fragment).
OS Thiobacillus cuprinus.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Comamonadaceae; Thiomonas.
OX NCBI_TaxID=36860;
RN [1]
RP SEQUENCE.
RC STRAIN=DSM 5494;
RA Marin I., Amaro A.M., Jerez C.A., Amils R., Abad J.P.;
RL Submitted (SEP-1995) to the SWISS-PROT data bank.
CC -1- MISCELLANEOUS: FOUND SPECIFICALLY IN CELLS CULTURED
CC CHEMOLITHOTROPHICALLY.
FT NON_TER 8 8
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 785 MW; 91487B06DDC2D76D CRC64;

Query Match 30.0%; Score 15; DB 1; Length 8;
Best Local Similarity 75.0%; Pred. No. 1.3e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 MAQE 6
DB 3 VAQE 6
I : : I I I

RESULT 4
FIBB_PAPAN STANDARD; PRT; 9 AA.
AC P19344;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).
GN FGB.
OS Papio anubis (Olive baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Papio.
OX NCBI_TaxID=9555;
RN [1]
RP SEQUENCE.
RC MEDLINE=84161822; PubMed=6423621;
RA Nakamura S., Takenaka O., Takahashi K.;
RT "Fibrinopeptides A and B of baboons (Papio anubis, Papio hamadryas,
RT and Theropithecus gelada): their amino acid sequences and
RT evolutionary rates and a molecular phylogeny for the baboons.";
RL J. Biochem. 94:1973-1978(1983).
CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC AGGREGATION.
CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
DR PIR: E28854; E28854.
DR InterPro: IPR002181; Fibrinogen_C.
DR PROSITE: PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
KW Blood coagulation; Plasma.
FT PEPTIDE 1 9 FIBRINOPEPTIDE B.
FT NON_TER 9 9
SQ SEQUENCE 9 AA; 1057 MW; DDFE71E9C7287B06 CRC64;

Query Match 28.0%; Score 14; DB 1; Length 9;
Best Local Similarity 75.0%; Pred. No. 1.3e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 QEAL 8
DB 2 QEGL 5
I : : I I I

RESULT 5
FIBB_PAPHA STANDARD; PRT; 9 AA.
AC P19343;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).
GN FGB.
OS Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Papio.
OX NCBI_TaxID=9557;
RN [1]
RP SEQUENCE.
RC MEDLINE=84161822; PubMed=6423621;
RA Nakamura S., Takenaka O., Takahashi K.;
RT "Fibrinopeptides A and B of baboons (Papio anubis, Papio hamadryas,
RT and Theropithecus gelada): their amino acid sequences and
RT evolutionary rates and a molecular phylogeny for the baboons.";
RL J. Biochem. 94:1973-1978(1983).
CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC AGGREGATION.
CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
DR PIR: E28854; E28854.
DR InterPro: IPR002181; Fibrinogen_C.
DR PROSITE: PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
KW Blood coagulation; Plasma.
FT PEPTIDE 1 9 FIBRINOPEPTIDE B.
FT NON_TER 9 9
SQ SEQUENCE 9 AA; 1057 MW; DDFE71E9C7287B06 CRC64;

Query Match 28.0%; Score 14; DB 1; Length 9;
Best Local Similarity 75.0%; Pred. No. 1.3e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 QEAL 8
DB 2 QEGL 5
I : : I I I

```

```

CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC AGGREGATION.
CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
DR PIR: D28854; D28854.
DR InterPro: IPR002181; Fibrinogen_C.
DR PROSITE: PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
KW Blood coagulation; Plasma.
FT PEPTIDE 1 9 FIBRINOPEPTIDE B.
FT NON_TER 9 9
SQ SEQUENCE 9 AA; 1076 MW; DDFE6409C7287B06 CRC64;

Query Match 28.0%; Score 14; DB 1; Length 9;
Best Local Similarity 75.0%; Pred. No. 1.3e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 QEAL 8
DB 2 QEGL 5
I : : I I I

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:04:04 ; Search time 23 Seconds  
(without alignments)  
22.491 Million cell updates/sec

Title: US-09-807-512-11  
Perfect score: 50  
Sequence: 1 MLMQAQALAF 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues  
Total number of hits satisfying chosen parameters: 441

Minimum DB seq length: 0  
Maximum DB seq length: 11

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	18	36.0	7	1	ALL2_CARMA	P81805 carcinus ma
2	16	32.0	9	1	RT33_BOVIN	P82926 bos taurus
3	15	30.0	8	1	CLP_THICU	P80488 thioabacilli
4	14	28.0	9	1	FIBB_PAPAN	P19344 papio anubi
5	14	28.0	9	1	FIBB_PAPHA	P19343 papio hamad
6	14	28.0	9	1	FIBB_THEGE	P19342 theropithec
7	14	28.0	9	1	MOSF_CLYJA	P19853 clypeaster
8	14	28.0	10	1	QZOG_COMTE	P80466 comamonas t
9	14	28.0	10	1	RL16_ACHLA	P29221 acholeplasm
10	13	26.0	7	1	ALL3_CARMA	P81806 carcinus ma
11	13	26.0	8	1	COXG_RAT	P80430 rattus norv
12	13	26.0	9	1	AL10_CARMA	P81813 carcinus ma
13	13	26.0	9	1	FAR2_CALVO	P41857 calliphora
14	13	26.0	9	1	FAR3_CALVO	P41858 calliphora
15	13	26.0	9	1	FRL1_SARBU	P83350 sarcophaga
16	13	26.0	9	1	ULAH_HUMAN	P31934 homo sapien
17	13	26.0	10	1	PSBF_CAPAN	Q03367 capsicum an
18	13	26.0	10	1	URAG_HUMAN	P32080 homo sapien
19	13	26.0	11	1	ULAG_HUMAN	P31933 homo sapien
20	12	24.0	7	1	FAR2_ASCSU	P31890 ascaris suu
21	12	24.0	9	1	FAR1_CALVO	P41856 calliphora
22	12	24.0	9	1	FAR4_CALVO	P41859 calliphora
23	12	24.0	9	1	FAR5_CALVO	P41860 calliphora
24	12	24.0	9	1	FAR6_CALVO	P41861 calliphora
25	12	24.0	9	1	FAR7_CALVO	P41862 calliphora
26	12	24.0	9	1	FIBB_MACFU	P19345 macaca fusc
27	12	24.0	10	1	FARP_MYTED	P42560 mytilus edu
28	12	24.0	10	1	TRP5_LEUMA	P81737 leucophaea
29	12	24.0	11	1	TKN_ELEMO	P01293 eledone mos
30	11	22.0	6	1	ACPH_RABIT	P25154 cryctolagus
31	11	22.0	6	1	TRPL_PSEPU	P36414 pseudomonas
32	11	22.0	7	1	CGF1_ENTFA	P20104 enterococcu
33	11	22.0	7	1	UF03_MOUSE	P38641 mus musculu

34	11	22.0	8	1	FAR8_CALVO	P41863 calliphora
35	11	22.0	8	1	NS3_MYCTU	P81152 mycobacteri
36	11	22.0	8	1	UPA1_HUMAN	P30087 homo sapien
37	11	22.0	9	1	FAR3_PENMO	P83318 penaeus mon
38	11	22.0	9	1	FARP_CALSI	P38495 callinectes
39	11	22.0	9	1	FIBB_ERYPA	P19346 erythrocebu
40	11	22.0	9	1	HUTU_KLEAE	P12381 klebsiella
41	11	22.0	9	1	SAMP_MUSCA	P19095 mustelus ca
42	11	22.0	10	1	ESL_LACCA	P81758 lactobacill
43	11	22.0	10	1	FARP_MANSE	P18523 manduca sex
44	11	22.0	10	1	NO40_TOBAC	P55962 nicotiana t
45	11	22.0	10	1	QZOB_COMTE	P80465 comamonas t

ALIGNMENTS

RESULT 1  
ALL2\_CARMA  
ID ALL2\_CARMA STANDARD; PRT; 7 AA.  
AC P81805;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Carcinustatin 2.  
OS Carcinus maenas (Common shore crab) (Green crab).  
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
OC Eubranchyura; Portunoidae; Fortuinidae; Carcinus.  
OX NCBI\_TaxID=6759;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;  
RX MEDLINE=98121193; PubMed=9461295;  
RA Duve H., Johnsen A.H., Maestro J.L., Scott A.G., Jaros P.P., Thorpe A.;  
RT "Isolation and identification of multiple neuropeptides of the allatostatin superfamily in the shore crab Carcinus maenas.";  
RL Eur. J. Biochem. 250:727-734(1997).  
CC -|- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.  
CC -|- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
KW Neuropeptide; Amidation; Multigene family.  
FT MOD\_RES 7  
SQ SEQUENCE 7 AA; 770 MW; 672879CDB5DDB70 CRC64;  
Query Match 36.0%; Score 18; DB 1; Length 7;  
Best Local Similarity 80.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	6	EALAF	10
Db	1	EAYAF	5

RESULT 2  
RT33\_BOVIN  
ID RT33\_BOVIN STANDARD; PRT; 9 AA.  
AC P82926;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Mitochondrial 28S ribosomal protein S33 (S33mt) (MRP-S33) (Fragment).  
GN MRP33.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Liver;  
RX MEDLINE=21276436; PubMed=11279123;  
RA Koc E.C., Burkhart W., Blackburn K., Moseley A., Spremull L.L.;

Query Match 28.0%; Score 14; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1e+04;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 QEA 7  
|||  
Db 5 QEA 7

## RESULT 15

A60722  
cryptic fimbrial protein - Serratia marcescens (strain US5) (fragment)  
C:Species: Serratia marcescens  
C:Date: 28-Apr-1993 #sequence\_revision 28-Apr-1993 #text\_change 17-Mar-1999  
C:Accession: A60722  
R:Moriya, T.; Kawabata, S.I.; Mizunoe, Y.; Amako, K.  
J. Bacteriol. 171, 6629-6636, 1989  
A:Title: A cryptic fimbrial gene in Serratia marcescens.  
A:Reference number: A60722; MUID:90078108; PMID:2574174  
A:Accession: A60722  
A:Molecule type: protein  
A:Residues: 1-10 <MOR>  
C:Comment: This protein was purified after expression of its gene in E. coli. and appeared

Query Match 28.0%; Score 14; DB 2; Length 10;  
Best Local Similarity 75.0%; Pred. No. 1e+04;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 EALA 9  
:|||  
Db 6 QALA 9

Search completed: October 7, 2003, 14:16:40  
Job time : 40 secs

Query Match		28.0%;	Score 14;	DB 2;	Length 8;	Best Local Similarity		75.0%;	Pred. No. 2.8e+05;	Matches 3;		Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Matches		3;		Conservative		0;		Mismatches		2;		Indels		0;	
Qy	6	EALAFI 11													
Db	2	EILDFI 7													
RESULT 9															
D28854															
fibrinopeptide B - olive baboon															
C:Species: Papio anubis, Papio hamadryas anubis (olive baboon)															
C:Date: 19-May-1989 #sequence_revision 19-May-1989 #text_change 18-Aug-2000															
C:Accession: D28854															
R:Nakamura, S.; Takenaka, O.; Takahashi, K.															
J. Biochem. 94, 1973-1978, 1983															
A:Title: Fibrinopeptides A and B of baboons (Papio anubis, Papio hamadryas, and Theropit															
A:Reference number: A91973; MUID:84161822; PMID:6423621															
A:Accession: D28854															
A:Molecule type: protein															
A:Residues: 1-9 <NAK>															
C:Superfamily: fibrinogen beta chain; fibrinogen beta/gamma homology; fibrinogen disulfid															
Query Match		28.0%;	Score 14;	DB 2;	Length 9;	Best Local Similarity		75.0%;	Pred. No. 2.8e+05;	Matches 3;		Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Matches		3;		Conservative		0;		Mismatches		1;		Indels		0;	
Qy	5	QEAL 8													
Db	2	QEGL 5													
RESULT 10															
E28854															
fibrinopeptide B - hamadryas baboon															
C:Species: Papio hamadryas (hamadryas baboon)															
C:Date: 19-May-1989 #sequence_revision 19-May-1989 #text_change 18-Aug-2000															
C:Accession: E28854															
R:Nakamura, S.; Takenaka, O.; Takahashi, K.															
J. Biochem. 94, 1973-1978, 1983															
A:Title: Fibrinopeptides A and B of baboons (Papio anubis, Papio hamadryas, and Theropit															
A:Reference number: A91973; MUID:84161822; PMID:6423621															
A:Accession: E28854															
A:Molecule type: protein															
A:Residues: 1-9 <NAK>															
C:Superfamily: fibrinogen beta chain; fibrinogen beta/gamma homology; fibrinogen disulfid															
Query Match		28.0%;	Score 14;	DB 2;	Length 9;	Best Local Similarity		75.0%;	Pred. No. 2.8e+05;	Matches 3;		Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Matches		3;		Conservative		0;		Mismatches		1;		Indels		0;	
Qy	5	QEAL 8													
Db	2	QEGL 5													
RESULT 11															
F28854															
fibrinopeptide B - gelada baboon															
C:Species: Theropithecus gelada (gelada baboon)															
C:Date: 19-May-1989 #sequence_revision 19-May-1989 #text_change 18-Aug-2000															
C:Accession: F28854															
R:Nakamura, S.; Takenaka, O.; Takahashi, K.															
J. Biochem. 94, 1973-1978, 1983															
A:Title: Fibrinopeptides A and B of baboons (Papio anubis, Papio hamadryas, and Theropit															
A:Reference number: A91973; MUID:84161822; PMID:6423621															
A:Accession: F28854															
A:Molecule type: protein															
A:Residues: 1-9 <NAK>															
C:Superfamily: fibrinogen beta chain; fibrinogen beta/gamma homology; fibrinogen disulfid															
Query Match		28.0%;	Score 14;	DB 2;	Length 9;	Best Local Similarity		75.0%;	Pred. No. 2.8e+05;	Matches 3;		Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Matches		3;		Conservative		0;		Mismatches		1;		Indels		0;	
Qy	5	QEAL 8													
Db	2	QEGL 5													
RESULT 12															
S66608															
quinoline 2-oxidoeductase gamma chain - Comamonas testosteroni (fragment)															
C:Species: Comamonas testosteroni															
C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999															
C:Accession: S66608															
R:Schach, S.; Tshisuaka, B.; Fetzner, S.; Lingens, F.															
Eur. J. Biochem. 232, 536-544, 1995															
A:Title: Quinoline 2-oxidoeductase and 2-oxo-1,2-dihydroquinoline 5,6-dioxygenase f															
A:Reference number: S66606; MUID:96035889; PMID:7556204															
A:Accession: S66608															
A:Molecule type: protein															
A:Residues: 1-9 <SCH>															
A:Experimental source: strain 63															
Query Match		28.0%;	Score 14;	DB 2;	Length 9;	Best Local Similarity		33.3%;	Pred. No. 2.8e+05;	Matches 2;		Conservative 3;	Mismatches 1;	Indels 0;	Gaps 0;
Matches		2;		Conservative		3;		Mismatches		1;		Indels		0;	
Qy	1	MLMAQE 6													
Db	1	MIQAEK 6													
RESULT 13															
JN0027															
[phe-6]-mosact - sea urchin (Clypeaster japonicus)															
C:Species: Clypeaster japonicus															
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Dec-1992															
C:Accession: JN0027															
R:Suzuki, N.; Kurita, M.; Yoshino, K.I.; Kajlura, H.; Nomura, K.; Yamaguchi, M.															
Zool. Sci. 4, 649-656, 1987															
A:Title: Purification and structure of mosact and its derivatives from the egg jelly															
A:Reference number: JN0025															
A:Accession: JN0027															
A:Molecule type: protein															
A:Residues: 1-9 <SUZ>															
C:Comment: Mosact, one of several sperm activating peptides located in egg jelly, and															
Query Match		28.0%;	Score 14;	DB 2;	Length 9;	Best Local Similarity		100.0%;	Pred. No. 2.8e+05;	Matches 3;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Matches		3;		Conservative		0;		Mismatches		0;		Indels		0;	
Qy	9	AFL 11													
Db	5	AFL 7													
RESULT 14															
S24190															
trypsin (EC 3.4.21.59) - bovine (fragment)															
C:Species: Bos primigenius taurus (cattle)															
C:Date: 19-Mar-1997 #sequence_revision 18-Aug-2000 #text_change 18-Aug-2000															
C:Accession: S24190															
R:Floriucci, L.; Erba, F.; Ascoli, F.															
Biol. Chem. Hoppe-Seyler 373, 483-490, 1992															
A:Title: Bovine trypsin: purification and characterization.															
A:Reference number: S24190; MUID:92384956; PMID:1515079															
A:Accession: S24190															
A:Status: preliminary															
A:Molecule type: protein															
A:Residues: 1-10 <FIO>															
C:Superfamily: trypsin; trypsin homology															
C:Keywords: hydrolase; serine proteinase; zymogen															

calsequestrin, cardiac and slow skeletal muscle - northern leopard frog (fragment)  
 N;Alternate names: 58K dihydropyridine-binding protein; aspartactin; calmitine; laminin-  
 C;Species: Rana pipiens (northern leopard frog)  
 C;Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 16-Jul-1999  
 C;Accession: A61230  
 R;McLeod, A.G.; Shen, A.C.Y.; Campbell, K.P.; Michalak, M.; Jorgensen, A.O.  
 Circ. Res. 69, 344-359, 1991  
 A;Title: Frog cardiac calsequestrin. Identification, characterization, and subcellular d  
 rdtum.  
 A;Reference number: A61230; MUID:91316784; PMID:1860177  
 A;Accession: A61230  
 A;Molecule type: protein  
 A;Residues: 1-19 <MC>  
 C;Comment: Calsequestrin is a high-capacity and moderate-affinity calcium binding protei  
 C;Comment: Calsequestrin acts as a calcium buffer, and the release of calcium bound to c  
 C;Comment: The fast skeletal muscle isoform of calsequestrin can be phosphorylated in vi  
 C;Superfamily: calsequestrin  
 C;Keywords: calcium binding; cardiac muscle; glycoprotein; heart; phosphoprotein; skelet  
 Query Match 30.0%; Score 15; DB 2; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 5 QEALAF 10  
 :|||  
 Db 1 EEGLNF 6

RESULT 4  
 S70721  
 heat shock protein C62.5 homolog - Salmonella typhimurium (fragment)  
 N;Alternate names: high temperature protein G  
 C;Species: Salmonella typhimurium  
 C;Date: 11-Mar-1998 #sequence\_revision 17-Apr-1998 #text\_change 17-Apr-1998  
 C;Accession: S70721  
 R;Ol, S.Y.; Li, Y.; Szyroki, A.; Giles, I.G.; Moir, A.; O'Connor, C.D.  
 Mol. Microbiol. 17, 523-531, 1995  
 A;Title: Salmonella typhimurium responses to a bactericidal protein from human neutroph  
 A;Reference number: S70719; MUID:96100451; PMID:8559071  
 A;Accession: S70721  
 A;Molecule type: protein  
 A;Residues: 1-10 <QIS>  
 A;Experimental source: strain SL1344  
 C;Keywords: ATP binding; heat shock; molecular chaperone  
 Query Match 30.0%; Score 15; DB 2; Length 10;  
 Best Local Similarity 33.3%; Pred. No. 6.3e+03;  
 Matches 3; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
 QY 2 LMAQEALAF 10  
 :|||  
 Db 1 MRGQETRGF 9

RESULT 5  
 PQ0682  
 photosystem I 17.5K D2 chain - common tobacco (fragment)  
 C;Species: Nicotiana tabacum (common tobacco)  
 C;Date: 19-May-1994 #sequence\_revision 19-May-1994 #text\_change 17-Mar-1999  
 C;Accession: PQ0682  
 R;Obokata, J.; Mikami, K.; Hayashida, N.; Nakamura, M.; Suglura, M.  
 Plant Physiol. 102, 1259-1267, 1993  
 A;Title: Molecular heterogeneity of photosystem I. psaD, psaE, psaF, psaH and psaL are a  
 A;Reference number: PQ0667; MUID:94105345; PMID:8278548  
 A;Accession: PQ0682  
 A;Molecule type: protein  
 A;Residues: 1-11 <OBO>  
 C;Superfamily: photosystem I chain II  
 C;Keywords: chloroplast; photosynthesis; photosystem I; thylakoid  
 Query Match 30.0%; Score 15; DB 2; Length 11;  
 Best Local Similarity 75.0%; Pred. No. 6.9e+03;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 AQEA 7  
 :|||  
 Db 1 AEEA 4

## RESULT 6

H54346  
 pyruvate synthase (EC 1.2.7.1) alpha chain - Pyrococcus furiosus (fragment)  
 C;Species: Pyrococcus furiosus  
 C;Date: 05-Jan-1996 #sequence\_revision 05-Jan-1996 #text\_change 05-May-2000  
 C;Accession: H54346  
 R;Blamey, J.M.; Adams, M.W.  
 Biochemistry 33, 1000-1007, 1994  
 A;Title: Characterization of an ancestral type of pyruvate ferredoxin oxidoreductase  
 A;Reference number: A54346; MUID:94137707; PMID:8305426  
 A;Accession: H54346  
 A;Status: preliminary  
 A;Molecule type: protein  
 A;Residues: 1-11 <BLA>  
 C;Keywords: coenzyme A; oxidoreductase

Query Match 30.0%; Score 15; DB 2; Length 11;  
 Best Local Similarity 30.0%; Pred. No. 6.9e+03;  
 Matches 3; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 MLMAQEALAF 10  
 :|||  
 Db 2 VMKGNEAAAW 11

## RESULT 7

A37765  
 hypothetical protein (csma 5' region) - Chloroflexus aurantiacus (fragment)  
 C;Species: Chloroflexus aurantiacus  
 C;Date: 31-May-1991 #sequence\_revision 31-May-1991 #text\_change 30-Sep-1993  
 C;Accession: A37765  
 R;Theroux, S.J.; Redlinger, T.E.; Fuller, R.C.; Robinson, S.J.  
 J. Bacteriol. 172, 4497-4504, 1990  
 A;Title: Gene encoding the 5.7-kilodalton chlorosome protein of Chloroflexus aurantia  
 A;Reference number: A37765; MUID:90330558; PMID:2376566  
 A;Accession: A37765  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-6 <THE>  
 A;Cross-references: GB:M33964

Query Match 28.0%; Score 14; DB 2; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 MAQ 5  
 :|||  
 Db 3 MAQ 5

## RESULT 8

S66296  
 Na<sup>+</sup>-transporting ATP synthase (EC 3.6.1.-) chain c - Acetobacterium woodii (fragment)  
 N;Alternate names: ATPase chain c  
 C;Species: Acetobacterium woodii  
 C;Date: 19-Mar-1997 #sequence\_revision 06-Jun-1997 #text\_change 07-May-1999  
 C;Accession: S66296  
 R;Reidlinger, J.; Mueller, V.  
 Eur. J. Biochem. 223, 275-283, 1994  
 A;Title: Purification of ATP synthase from Acetobacterium woodii and identification a  
 A;Reference number: S45648; MUID:94307271; PMID:8033902  
 A;Accession: S66296  
 A;Molecule type: protein  
 A;Residues: 1-8 <REI>  
 A;Experimental source: DSM 1030  
 C;Keywords: hydrolase

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:11:09 ; Search time 38 Seconds  
(without alignments)  
27.838 Million cell updates/sec

Title: US-09-807-512-11

Perfect score: 50

Sequence: 1 MLMQAQALAF 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 segs, 96168682 residues

Total number of hits satisfying chosen parameters: 1326

Minimum DB seq length: 0

Maximum DB seq length: 11

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_76:\*

1: p1r1.\*

2: p1r2.\*

3: p1r3.\*

4: p1r4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	18	36.0	9	2	A33527 fructose-2,6-bisph
2	17	34.0	11	2	A35594 buccalin - Califor
3	15	30.0	9	2	A61230 calsequestrin, car
4	15	30.0	10	2	S70721 heat shock protein
5	15	30.0	11	2	P00682 photosystem I 17.5
6	15	30.0	11	2	H54346 pyruvate synthase
7	14	28.0	16	2	A37765 hypothetical prote
8	14	28.0	8	2	S66296 Na+-transporting A
9	14	28.0	9	2	D28854 fibrinopeptide B -
10	14	28.0	9	2	E28854 fibrinopeptide B -
11	14	28.0	9	2	F28854 fibrinopeptide B -
12	14	28.0	9	2	S66608 quinoline 2-oxidor
13	14	28.0	9	2	JN0027 [Phe-6]-mosact - s
14	14	28.0	10	2	S24190 triptase (EC 3.4.2
15	14	28.0	10	2	A60722 cryptic fibrial p
16	14	28.0	10	2	F41839 ribosomal protein
17	14	28.0	10	2	PH0113 alpha-amylase (EC
18	14	28.0	11	2	B50769 Ig H2 chain - Paci
19	13	26.0	7	2	P02466 Ig heavy chain CRD
20	13	26.0	8	2	S66646 cardioacceleratory
21	13	26.0	8	2	S55381 cytochrome-c oxida
22	13	26.0	9	2	B41978 callifMRamide 2 -
23	13	26.0	9	2	C41978 callifMRamide 3 -
24	13	26.0	10	2	S28055 cytochrome b559 co
25	13	26.0	10	2	I44644 neurotoxin-associ
26	13	26.0	10	4	SI4943 UGA3 leader peptid
27	13	26.0	11	2	S78026 ribosomal protein
28	13	26.0	11	2	PH0924 T-cell receptor be
29	13	26.0	11	4	I52708 ELAV-like neuronal

30	13	26.0	11	4	S52252 hypothetical prote
31	12	24.0	4	2	S18401 thyroglobulin - do
32	12	24.0	6	2	T11779 phosphoglycerate t
33	12	24.0	6	4	S15596 orf 3 rara 5'-regl
34	12	24.0	7	2	S55548 mcrB protein - Esc
35	12	24.0	7	2	P80254 18K protein 5507 -
36	12	24.0	7	2	P00777 NADH2 dehydrogenas
37	12	24.0	8	2	A37521 R-phycocerythrin ga
38	12	24.0	8	2	T10952 hypothetical prote
39	12	24.0	8	2	I48935 apolipoprotein A-I
40	12	24.0	8	2	B47594 aspartate kinase (
41	12	24.0	8	2	S53008 citrate synthase -
42	12	24.0	9	2	C24180 fibrinogen beta ch
43	12	24.0	9	2	S36898 ribosomal protein
44	12	24.0	9	2	A41978 callifMRamide 1 -
45	12	24.0	9	2	D41978 callifMRamide 4 -

#### ALIGNMENTS

##### RESULT 1

A33527

fructose-2,6-bisphosphate 2-phosphatase (EC 3.1.3.46) - rat (fragment)

N:Alternate names: fructose-2,6-bisphosphatase

C:Species: Rattus norvegicus (Norway rat)

C:Date: 27-Feb-1990 #sequence\_revision 27-Feb-1990 #text\_change 23-Jun-1993

C:Accession: A33527

R:Kitamura, K.; Uyeda, K.; Hartman, F.C.; Kangawa, K.; Matsuo, H.

J. Biol. Chem. 264, 6344-6348, 1989

A:Title: Catalytic site of rat liver and bovine heart fructose-6-phosphate,2-kinase:f

A:Reference number: A33527; MUID:89197937; PMID:2539378

A:Accession: A33527

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-9 <KIT>

C:Keywords: phosphoric monoester hydrolase

Query Match 36.0%; Score 18; DB 2; Length 9;

Best Local Similarity 71.4%; Pred. No. 2.8e+05;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 QEALAF 11

Db 2 QXALAL 8

##### RESULT 2

A35594

buccalin - California sea hare

C:Species: Aplysia californica (California sea hare)

C:Date: 14-Sep-1990 #sequence\_revision 14-Sep-1990 #text\_change 24-Jun-1993

C:Accession: A35594

R:Cropper, E.C.; Miller, M.W.; Tenenbaum, R.; Kolks, M.A.G.; Kupfermann, I.; Weiss, K

Proc. Natl. Acad. Sci. U.S.A. 85, 6177-6181, 1988

A:Title: Structure and action of buccalin: a modulatory neuropeptide localized to an

A:Reference number: A35594; MUID:88320404; PMID:3413086

A:Accession: A35594

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-11 <CRO>

Query Match 34.0%; Score 17; DB 2; Length 11;

Best Local Similarity 60.0%; Pred. No. 2.6e+03;

Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 6 EALAF 10

Db 3 DSLAF 7

##### RESULT 3

A61230

XX Human; single nucleotide polymorphism; SNP; paternity test;  
KW forensic test; aberrant protein expression.  
XX Homo sapiens.  
XX  
XX WO200151670-A2.  
XX  
XX 19-JUL-2001.  
XX  
XX 05-JAN-2001; 2001WO-US00322.  
XX  
XX 07-JAN-2000; 2000US-0174962.  
XX  
XX (CURA-) CURAGEN CORP.  
XX  
XX Shinkets RA, Leach MD;  
XX  
XX WPI: 2001-451871/48.  
XX N-PSDB; AAH89309.  
XX  
XX Isolated human polynucleotides containing single nucleotide  
PT polymorphisms, useful for the treatment and diagnosis of e.g. cancer,  
PT infection and diabetes -  
XX  
XX Disclosure; Page 317; 475pp; English.  
XX  
XX The present invention relates to human nucleic acids containing single  
CC nucleotide polymorphisms (SNPs). These can be used in forensic and  
CC paternity tests, and to aid in the treatment of diseases associated with  
CC aberrant protein expression, including cancer, amyloidosis, diabetes,  
CC Alzheimer's disease, Down's syndrome, oedema, lupus (SLE), vasculitis,  
CC glomerulonephritis, haemolytic anaemia, thrombocytopaenia, arthritis,  
CC meningitis, muscular disorders, dementia, neurological diseases, tubercu-  
CC sclerosis, male infertility, hypercalcaemia, blood pressure disorders,  
CC osteoporosis, pathogenic infections, hypercholesterolaemia, obesity or  
CC autoimmunity. The present sequence is a peptide encoded by a  
CC polymorphism-containing oligonucleotide fragment of the invention.  
XX  
SQ Sequence 11 AA;  
Query Match 48.0%; Score 24; DB 22; Length 11;  
Best Local Similarity 55.6%; Pred. No. 3.9e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
Oy 3 MAQELAF 11  
:|: | |  
Db 1 VAQDGLNFL 9

Search completed: October 7, 2003, 14:13:44  
Job time : 84 secs

```
XX SQ Sequence 10 AA;
Query Match 48.0%; Score 24; DB 17; Length 10;
Best Local Similarity 75.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 AQEALAF 11
Db 1:|||||
2 AKELAA 9

RESULT 13
AAU71255
ID AAU71255 standard; Peptide; 10 AA.
XX AC AAU71255;
XX DT 26-FEB-2002 (first entry)
XX DE Human MHC class I molecule HLA-A2 binding 103P3E8 peptide #62.
KW 103P3E8; prostate; bladder; kidney; colon; lung; breast; rectum; stomach;
KW tumour; cancer; cytostatic; gene therapy; antibody therapy; ribozyme;
KW single chain monoclonal antibody; serum; blood; urine; tissue; human;
KW chromosome 9q13-q21.
XX OS Homo sapiens.
XX PN WO200179557-A2.
XX PD 25-OCT-2001.
XX PF 12-APR-2001; 2001WO-US12181.
XX PR 12-APR-2000; 2000US-196647P.
XX PA (UROG-) UROGENESYS INC.
XX PI Faris M, Challita-Eid PM, Raitano AB, Mitchell SC, Afar DEH;
XX PI Jakobovits A;
XX DR WPI; 2002-061976/08.
XX PT Monitoring 103P3E8 gene products in sample from patient (suspected of)
XX PT having cancer, useful for diagnosing, managing or treating cancers,
XX PT e.g. prostate cancer, comprises determining presence of aberrant
XX PT 103P3E8 gene products
XX PS Disclosure; Page 85; 128pp; English.
XX CC Sequences AAU71093-AAU71796 represent the 103P3E8-related protein and
XX CC peptide fragments of the protein. 103P3E8 exhibits tissue specific
XX CC expression in normal adult tissue, but it is also aberrantly expressed in
XX CC many cancers including tumours of the prostate, bladder, kidney, colon,
XX CC lung, breast, rectum and stomach. The 103P3E8 polynucleotide, its related
XX CC protein and peptide fragments and specific PCR primers are therefore
XX CC useful for diagnosing and treating cancer. A vector comprising a
XX CC polynucleotide which encodes a single chain monoclonal antibody, that
XX CC immunospecifically binds to an 103P3E8-related protein, and a ribozyme
XX CC capable of cleaving a polynucleotide having the 103P3E8 coding sequence,
XX CC are both useful in the preparation of a composition for treating a
XX CC patient with a cancer that expresses 103P3E8. The sequences can be used
XX CC in diagnostic methods to monitor the level of 103P3E8 gene products in
XX CC serum, blood, urine and tissue and to thereby detect the presence of
XX CC cancerous cells.
XX SQ Sequence 10 AA;
Query Match 48.0%; Score 24; DB 23; Length 10;
Best Local Similarity 60.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 AQEALAF 11
Db 1:|||||
2 AKELAA 9

RESULT 14
AAU00199
ID AAU00199 standard; Peptide; 11 AA.
XX AC AAU00199;
XX DT 01-OCT-2001 (first entry)
XX DE Human protein fragment SEQ ID NO: 739.
KW Human; single nucleotide polymorphism; SNP; paternity test;
KW forensic test; aberrant protein expression.
XX OS Homo sapiens.
XX PN WO200151670-A2.
XX PD 19-JUL-2001.
XX PF 05-JAN-2001; 2001WO-US00322.
XX PR 07-JAN-2000; 2000US-0174962.
XX PA (CURA-) CURAGEN CORP.
XX PI Shimkets RA, Leach MD;
XX DR WPI; 2001-451871/48.
XX DR N-PSDB; AAH89308.
XX PT Isolated human polynucleotides containing single nucleotide
XX PT polymorphisms, useful for the treatment and diagnosis of e.g. cancer,
XX PT infection and diabetes
XX PS Disclosure; Page 317; 475pp; English.
XX CC The present invention relates to human nucleic acids containing single
XX CC nucleotide polymorphisms (SNPs). These can be used in forensic and
XX CC paternity tests, and to aid in the treatment of diseases associated with
XX CC aberrant protein expression, including cancer, amyloidosis, diabetes,
XX CC Alzheimer's disease, Down's syndrome, oedema, lupus (SLE), vasculitis,
XX CC glomerulonephritis, haemolytic anaemia, thrombocytopaenia, arthritis,
XX CC meningitis, muscular disorders, dementia, neurological diseases, tubercous
XX CC sclerosis, male infertility, hypercalcaemia, blood pressure disorders,
XX CC osteoporosis, pathogenic infections, hypercholesterolaemia, obesity or
XX CC autoimmunity. The present sequence is a peptide encoded by a
XX CC polymorphism-containing oligonucleotide fragment of the invention.
XX SQ Sequence 11 AA;
Query Match 48.0%; Score 24; DB 22; Length 11;
Best Local Similarity 55.6%; Pred. No. 3.9e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 MAQEALAF 11
Db 1 VAQDGLDF 9

RESULT 15
AAU00200
ID AAU00200 standard; Peptide; 11 AA.
XX AC AAU00200;
XX DT 01-OCT-2001 (first entry)
XX DE Human protein fragment SEQ ID NO: 740.
```



PA (CAMB-) CAMBRIDGE NEUROSCIENCE.  
 XX Gwynne DI, Marchionni M, Sklar R;  
 XX WPI; 1995-006353/01.  
 DR  
 XX  
 XX Treating mammalian muscle diseases and disorders - by admin. of  
 PT GGF2 and other specified polypeptide(s) which bind the p185erbB2  
 PT receptor.  
 XX  
 XX Example 9; Page 91; 241pp; English.  
 PS  
 XX AAR67153-R67163 and AAR67174-R67183 are bovine glial cell growth  
 CC factor-1 (GGF-1) peptide fragments. A 30-36kD polypeptide factor  
 CC known to have glial cell mitogenic activity, which includes one  
 CC of the above peptide fragments in its amino acid sequence is  
 CC claimed. The polypeptide can be used to treat a variety of  
 CC mammalian skeletal, cardiac and smooth muscle diseases,  
 CC including acetylcholine receptor deficiency.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 CC  
 XX Sequence 10 AA;  
 XX  
 XX Query Match 48.0%; Score 24; DB 16; Length 10;  
 XX Best Local Similarity 75.0%; Pred. No. 3.5e+02;  
 XX Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 AQEALAF 11  
 DB :|||||  
 2 AKEALAL 9  
 RESULT 11  
 AAW03313  
 ID AAW03313 standard; Peptide; 10 AA.  
 XX  
 AC AAW03313;  
 XX  
 DT 22-OCT-1996 (first entry)  
 XX  
 XX Tyrosinase peptide T10199 (aal99-208).  
 DE  
 XX Melanoma; tyrosinase; immunogen; vaccine; cancer; immunotherapy.  
 KW  
 XX Synthetic.  
 OS  
 XX WO9621734-A2.  
 PN  
 XX 18-JUL-1996.  
 PD  
 XX 11-JAN-1996; 96WO-US00473.  
 PF  
 XX 10-JAN-1995; 95US-0370909.  
 PR  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA  
 XX Robbins PF, Rosenberg SA;  
 PI  
 XX WPI; 1996-342287/34.  
 DR  
 XX Nucleic acid sequence encoding p15 melanoma antigen - and  
 XX immunogenic peptide(s) derived from it, useful for diagnosis,  
 PT prevention or treatment of melanoma  
 PT  
 XX Example 2; Page 64; 98pp; English.  
 PS  
 XX A series of synthetic peptides (AAW03307-15 plus AAW03304-05) are  
 CC based on putative peptide binding motifs to HLA-A24 in human  
 CC tyrosinase melanoma antigen (AAW03306). Peptide AAW10199 (AAW03313)  
 CC corresponds to amino acids 199-208 of tyrosinase. The peptides  
 CC were tested for their ability to sensitize HLA-A24+ 888  
 CC Epstein-Barr virus-transformed B-cells to tumour infiltrating  
 CC lymphocyte (TIL) lysis and their ability to render melanoma 888

CC cells to stimulate GM-CSF release from TIL. Positive results  
 CC were obtd. with peptides T9206 (AAW03304) and AAT10206 (AAW03305).  
 CC These peptides can be used in melanoma vaccines.  
 XX  
 XX Sequence 10 AA;  
 XX  
 XX Query Match 48.0%; Score 24; DB 17; Length 10;  
 XX Best Local Similarity 75.0%; Pred. No. 3.5e+02;  
 XX Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 4 AQEALAF 11  
 DB :|||||  
 3 AHEAPAF 10  
 RESULT 12  
 AAR86642  
 ID AAR86642 standard; peptide; 10 AA.  
 XX  
 AC AAR86642;  
 XX  
 DT 03-JUL-1996 (first entry)  
 XX  
 XX GGF I tryptic peptide GGF-I 14 #1.  
 DE  
 XX Glial growth factor; GGF; human; hGGF2; Schwann cell; mitogenesis; GGF-I;  
 KW glial cell; therapy; peripheral nerve damage; demyelination; bovine;  
 KW glial tumour; fibroblast proliferation; wound repair; multiple sclerosis;  
 KW neurodegenerative disorder; neural regeneration; acetylcholine receptor.  
 XX  
 OS Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FH Misc-difference 1 /label= Lys, Arg  
 FT  
 FT  
 XX WO9532724-A1.  
 PN  
 XX 07-DEC-1995.  
 XX  
 XX 25-MAY-1995; 95WO-US06846.  
 XX  
 XX 26-MAY-1994; 94US-0249322.  
 PR  
 XX (CAMB-) CAMBRIDGE NEUROSCIENCE.  
 PA  
 XX (LUDW-) LUDWIG INST CANCER RES.  
 XX  
 XX Chen MS, Goodearl AD, Hiles I, Marchionni M, Minghetti L;  
 PI Stroobant P, Waterfield M;  
 PI  
 XX WPI; 1996-030329/03.  
 DR  
 XX N-PSDB; AAT06646, AAT06647, AAT06667, AAT06688, AAT06690.  
 DR  
 XX Glial growth factors with Schwann cell mitogenic activity - for  
 PT prophylaxis or treatment of nerve disorders, e.g. Multiple Sclerosis  
 PT  
 XX Example 2; Fig 9; 199pp; English.  
 PS  
 XX AAR86629-R86657 represent fragments of bovine glial growth factor I  
 CC (GGF-I). These fragments were obtained by lysyl endopeptidase and  
 CC protease V8 digests. These sequences have Schwann cell mitogenic  
 CC activity in the presence of foetal calf plasma. These sequences (and  
 CC human GGF2, see AAR86628) are used to stimulate mitogenesis of glial  
 CC cells, for prophylaxis or treatment of a pathological condition of  
 CC the nervous system in a mammal. Also, for identification of a receptor,  
 CC for treatment of conditions of peripheral nerve damage (e.g.  
 CC demyelination/damage/loss of Schwann cells), treatment of  
 CC neurodegenerative disorders in mammals, for inducing neural  
 CC regeneration, fibroblast proliferation or wound repair. Glial tumours  
 CC can be treated by inhibiting this sequence binding to its receptor.  
 CC These peptides are also able to induce synthesis of acetylcholine  
 CC receptor, and is useful in prophylaxis or treatment of multiple  
 CC sclerosis in a patient.



CC cancer and stomach cancer including in the form of vaccines. The present  
 CC sequence is that of a tumour antigen peptide, useful to the invention.

SQ Sequence 9 AA;

Query Match 54.0%; Score 27; DB 23; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 9.3e+05;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LMAQEALA 9

I:|||||  
 . 1 LLAQEAAA 8

RESULT 6

AAR28520  
 ID AAR28520 standard; Protein; 9 AA.

XX AC AAR28520;

XX DT 25-MAR-2003 (updated)  
 DT 18-MAR-1993 (first entry)

XX XX Novel peptide GGF-I 14.

XX KW Glial growth factor-II; GGF-II; backtranslation; bovine cDNA library;  
 KW polymerase chain reaction; PCR; amplify; pituitary RNA; exons;  
 KW alternative splice pattern.

XX OS Synthetic.

XX PN W09218627-A1.

XX XX 29-OCT-1992.

XX PF 03-APR-1992; 92WO-GB00595.

XX PR 10-APR-1991; 91GB-0007566.

XX XX (CAMB-) CAMBRIDGE NEUROSCIENCE INC.  
 PA (LUDW-) LUDWIG INST CANCER RES.

XX PI Goodearl ADJ, Marchionni MA, Minghetti L, Stroobant P;  
 PI Waterfield MD;

XX DR WPI; 1992-382111/46.

XX PT Glial mitogenic polypeptide factors - used in diagnosis,  
 PT prophylaxis or treatment of patho-physiological conditions of the  
 PT nervous system

XX PS Disclosure; Fig 10; 98pp; English.

XX XX The sequences given in AAR28514-25 are peptides derived from bovine  
 CC glial growth factor-I (GGF-I). Backtranslation of these sequences may  
 CC provide probes and primers which can be used for the isolation of the  
 CC GGF-I cDNA. A bovine cDNA library may be probed with the  
 CC backtranslated probes. Isolated open reading frames may then be  
 CC amplified using polymerase chain reaction on pituitary RNA.  
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 9 AA;

Query Match 48.0%; Score 24; DB 13; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 9.3e+05;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 AQEALAF 11

I:|||||  
 1 AKEALAA 8

RESULT 7

XX KW Glial growth factor; GGF; heregulin; mitogenesis;

AAR46863

ID AAR46863 standard; Protein; 9 AA.

XX AC AAR46863;

XX DT 25-MAR-2003 (updated)  
 DT 10-MAR-2003 (updated)  
 DT 28-JUL-1994 (first entry)

XX XX Cell proliferation inhibitor.

XX KW Glial growth factor; GGF; heregulin; mitogenesis;  
 KW Schwann cell; tumour; central nervous system;  
 KW antiproliferative; inhibitor.

XX OS Unidentified.

XX PN W09403644-A1.

XX XX 17-FEB-1994.

XX PF 10-AUG-1993; 93WO-US07491.

XX PR 10-AUG-1992; 92US-0927337.

XX PR 23-SEP-1992; 92US-0951747.

XX PR 01-DEC-1992; 92US-0984085.

XX PR 29-JAN-1993; 93US-0011396.

XX PA (CAMB-) CAMBRIDGE NEUROSCIENCE INC.

XX PI Gwynne DI, Marchionni M, McBurney RN;

XX XX WPI; 1994-065731/08.

XX PT Glial growth factor DNA encoding numerous polypeptide factors  
 PT used for inhibiting cell proliferation - for treating carcinoma  
 PT and nervous disorders

XX PS Disclosure; Fig 10; 178pp; English.

XX CC Factors useful in the methods of the invention are  
 CC basic polypeptide factors having antiproliferative  
 CC activity when contacted with cells, including cells of the  
 CC nervous system and specifically Schwann cells, and contg.  
 CC within their amino acid sequences at least one of the polypeptides  
 CC given in AAR46857-68.  
 CC (Updated on 10-MAR-2003 to add missing OS field.)  
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 9 AA;

Query Match 48.0%; Score 24; DB 15; Length 9;

Best Local Similarity 75.0%; Pred. No. 9.3e+05;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 AQEALAF 11

I:|||||  
 1 AKEALAA 8

RESULT 8

AAR55670  
 ID AAR55670 standard; Protein; 9 AA.

XX AC AAR55670;

XX DT 25-MAR-2003 (updated)  
 DT 10-MAR-2003 (updated)  
 DT 28-JUL-1994 (first entry)

XX DE 30-36 kD glial growth factor polypeptide.

XX KW Glial growth factor; GGF; heregulin; mitogenesis;

PS Disclosure; Page 17; 61pp; English.

XX CC The invention relates to a composition (I) comprising an immune effector cell and a cell penetrating peptide (CPP) associated with an antigen or antibody. Also included are (1) a vaccine comprising (I), CPP associated with an antigen, and a pharmaceutically acceptable carrier (2) preparing a composition for a disease, by providing (I) and (2) associated with an antigen for disease, and introducing the antigen-associated CPP to (I), where antigen enters into the cell. The antigens are, for example, tumour antigen derived epitopes recognised by tumour infiltrating lymphocytes (TIL) of HLA (human leukocyte antigen) class I or II. The composition is useful for enhancing immunity in an animal to a disease, by administering a mature dendritic cell comprising CPP associated with an antigen to disease, to the animal, such that following the administration, animal is protected from disease, where the animal comprises both CD4+ and CD8+ T cells. It is also useful for treating a disease (e.g. cancer, tumour, melanoma, thymoma, lymphoma, sarcoma, lung cancer, non-Hodgkin's lymphoma, leukaemia, Hodgkin's lymphoma, uterine cancer, cervical cancer, bladder cancer, kidney cancer, adenocarcinoma, breast cancer, prostate cancer, ovarian cancer and pancreatic cancer). The animal is further subjected to a cancer treatment including surgery, radiation, chemotherapy or gene therapy. The administration of (I), preferably dendritic cell is prior to, subsequent to or concurrent with, the cancer treatment. The present sequence is a tumour antigen derived epitope for inclusion in the composition of the invention.

XX SQ Sequence 10 AA;

Query Match 92.0%; Score 46; DB 23; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.021;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLMQAEALAF 10  
|||||  
DB 1 MLMQAEALAF 10

RESULT 4  
AAY70856

ID AAY70856 standard; peptide; 10 AA.

XX AC AAY70856;

XX DT 31-JUL-2000 (first entry)

XX DE CTL epitope-2 of human CAMEL protein.

XX KW CAMEL; CTL-recognised Antigen on MELanoma; cytotoxic T lymphocyte; CTL;  
KW tumour-associated antigen; LAGE-1; NY-ESO-1; anticancer; melanoma;  
KW human; cancer; immunotherapy; immunogenic peptide; immune response.

XX OS Homo sapiens.

XX PN WO200023584-A1.

XX PD 27-APR-2000.

XX PF 15-OCT-1999; 99WO-EP07832.

XX PR 16-OCT-1998; 98EP-0119583.

XX PA (BOEH ) BOEHRINGER INGELHEIM INT GMBH.  
PA (UYHO-) UNIV HOSPITAL LEIDEN.

XX PI Schrier PI, Aarnoudse CA, Heider K, Klade C;

XX DR WPI; 2000-339685/29.

XX PT Tumor-associated antigen useful for cancer immunotherapy is encoded by  
PT the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA -

XX PS Claim 5; Page 34; 73pp; English.

XX

CC The present sequence is an immunogenic peptide of human tumour-associated antigen CAMEL (Cytotoxic T lymphocytes (CTL)-recognised Antigen on MELanoma). This peptide is a CTL epitope, that has the ability to elicit a CTL response. It corresponds to residues 2-11 of the CAMEL protein. CAMEL protein is encoded by the LAGE-1 gene, a tumour-specific antigen. It is different from the LAGE-1 protein, since it is translated from a different open reading frame (ORF-1). It shows strong homology with NY-ESO-1, a melanoma specific tumour antigen. The tumour-associated antigen displayed on melanoma cells is recognised by cytotoxic T lymphocytes. CAMEL is expressed in tumour cell lines, tumour tissues (e.g. breast and lung) and in restricted number of healthy tissues. This sequence has anticancer activity. CAMEL tumour antigen and immunogenic peptides derived from it are useful for cancer immunotherapy. They have the potential to induce an immune response, by eliciting a CTL response. The DNA molecule is used to construct recombinant or fusion proteins.

XX SQ Sequence 10 AA;

Query Match 90.0%; Score 45; DB 21; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.033;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LMAQEALAF 11  
|||||  
DB 1 LMAQEALAF 10

RESULT 5

ABB97090

ID ABB97090 standard; Peptide; 9 AA.

XX AC ABB97090;

XX DT 21-JUN-2002 (first entry)

XX DE Human tumour antigen related peptide 141.

XX KW Human; tumour; antigen; HLA-A2; cytotoxic T cell; cytostatic; cancer;  
KW vaccine.

XX OS Homo sapiens.

XX PN WO200210369-A1.

XX PD 07-FEB-2002.

XX PF 30-JUL-2001; 2001WO-JP06526.

XX PR 31-JUL-2000; 2000JP-0231814.

XX PA (ITOH/) ITOH K.

XX PI Itoh K;

XX PN WPI; 2002-291857/33.

XX PT Tumor antigens inducing and/or activating HLA-A2-restricted  
PT tumor-specific cytotoxic T cells, useful in diagnosis of and screening  
PT drugs e.g. cancer vaccines for specific treatment of pancreatic cancer

XX PS Example 4; Fig 10; 127pp; Japanese.

XX CC The invention relates to a peptide comprising an amino acid sequence selected from 44 fully defined amino acid sequences (ABB96906-ABB969549) and a polypeptide comprising an amino acid sequence selected from the 9 fully defined amino acid sequences (ABB97143-ABB97151). The above comprise a tumour antigen inducing or activating HLA-A2-restricted tumour-specific cytotoxic T cells, which recognise HLA-A2 and a tumour antigen peptide and is thus activated. The peptides and polypeptides have cytostatic activity. The tumour antigen is useful in diagnosis of and screening drugs for specific treatment of pancreatic cancer, colon

PT the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA -

PS Claim 4; Page 34; 73pp; English.

XX The present sequence is an immunogenic peptide of human tumour-associated

CC antigen CAMEL (Cytotoxic T lymphocytes (CTL)-recognised Antigen on

CC MELANOMA). This peptide is a CTL epitope, that has the ability to elicit

CC a CTL response. It corresponds to residues 1-11 of the CAMEL protein.

CC CAMEL protein is encoded by the LAGE-1 gene, a tumour-specific antigen.

CC It is different from the LAGE-1 protein, since it is translated from a

CC different open reading frame (ORF-1). It shows strong homology with

CC NY-ESO-1, a melanoma specific tumour antigen. The tumour-associated

CC antigen displayed on melanoma cells is recognised by cytotoxic T

CC lymphocytes. CAMEL is expressed in tumour cell lines, tumour tissues

CC (e.g. breast and lung) and in restricted number of healthy tissues. This

CC sequence has anticancer activity. CAMEL tumour antigen and immunogenic

CC peptides derived from it are useful for cancer immunotherapy. They have

CC the potential to induce an immune response, by eliciting a CTL response.

CC The DNA molecule is used to construct recombinant or fusion proteins.

XX

SQ Sequence 11 AA;

Query Match 100.0%; Score 50; DB 21; Length 11;

Best Local Similarity 100.0%; Pred. No. 0.0039;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLMAQEALAF 11

Db 1 MLMAQEALAF 11

RESULT 2

ABG66804

ID ABG66804 standard; Peptide; 11 AA.

XX

AC ABG66804;

XX

DT 24-SEP-2002 (first entry)

XX

DE Tumour antigen LAGE/CAMEL, HLA-A2 epitope.

XX

KW Beta-2 microglobulin; beta-2m; cytotoxic T lymphocyte; CTL; HLA;

KW human leukocyte antigen; fusion protein; epitope; cytostatic; tumour;

KW gastrointestinal tumour; colorectal cancer; gastro-oesophageal cancer;

KW liver cancer; biliary tract cancer; pancreatic cancer; vaccine;

KW prostatic cancer; testicular cancer; lung cancer; breast cancer;

KW malignant melanoma; mesothelioma; brain tumour; ovarian cancer;

KW uterine cancer; cervical cancer; head and neck cancer; bladder cancer;

KW Kaposi's sarcoma; renal carcinoma; leukaemia; lymphoma;

KW acquired immunodeficiency syndrome; AIDS-related lymphoma.

XX

OS Homo sapiens.

XX

PN WO200236146-A2.

XX

PD 10-MAY-2002.

XX

PF 01-NOV-2001; 2001WO-GB04844.

XX

PR 02-NOV-2000; 2000GB-0026812.

XX

XX (ISIS-) ISIS INNOVATION LTD.

PA

XX Tafuro S, Meier U, McMichael AJ, Bell JI, Layton G, Hunter M;

PI WPI; 2002-508108/54.

XX

DR New polynucleotide capable of expressing an epitope-beta2m fusion

XX protein useful for generating cytotoxic T lymphocyte responses against

PT tumour and in restoring antigen presentation in the tumour of a host

PT

XX Disclosure; Page 25; 46pp; English.

PS

XX The invention relates to a new polynucleotide capable of expressing an

CC epitope-beta-2m fusion protein useful for generating cytotoxic T

CC lymphocyte (CTL) responses against a tumour or in restoring antigen

CC presentation in the tumour of a host. Also included are a polynucleotide

CC capable of expressing an epitope-beta-2m fusion protein in combination

CC with a vaccination agent that stimulates a CTL response against the

CC epitope of the fusion protein for simultaneous, separate or sequential

CC use in the treatment of cancer and a method of treating a tumour by

CC administering a capable of expressing an epitope-beta-2m fusion protein,

CC and optionally a vaccination agent that stimulates a CTL response against

CC the epitope of the fusion protein. The polynucleotide is useful for

CC generating CTL responses against tumours, for restoring antigen

CC presentation in the tumour, and subsequently for treating cancers, such

CC as gastrointestinal tumour, prostatic, testicular, lung or breast cancer,

CC malignant melanoma, mesothelioma, brain tumour, ovarian cancer, uterine

CC cancer including cervical cancer, cancer of the head and neck, bladder

CC cancer, Kaposi's sarcoma, AIDS (acquired immunodeficiency syndrome)-

CC related Kaposi's sarcoma, sarcomas, osteosarcoma, renal carcinoma, and

CC haematopoietic malignant tumours such as leukaemia and lymphoma.

CC The epitope is an HLA (human leukocyte antigen) peptide derived from a

CC viral or tumour antigen. The present sequence is a tumour HLA epitope

CC used in the fusion proteins of the invention.

XX

SQ Sequence 11 AA;

Query Match 100.0%; Score 50; DB 23; Length 11;

Best Local Similarity 100.0%; Pred. No. 0.0039;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLMAQEALAF 11

Db 1 MLMAQEALAF 11

RESULT 3

ABG79072

ID ABG79072 standard; Peptide; 10 AA.

XX

AC ABG79072;

XX

DT 15-NOV-2002 (first entry)

XX

DE Human CAMEL class I HLA widely expressed antigen peptide #1.

XX

KW Cell penetrating peptide; cancer; tumour; melanoma; thymoma; antigen;

KW lymphoma; sarcoma; lung cancer; non-Hodgkin's lymphoma; leukaemia;

KW Hodgkin's lymphoma; uterine cancer; cervical cancer; bladder cancer;

KW kidney cancer; adenocarcinoma; breast cancer; prostate cancer;

KW ovarian cancer; pancreatic cancer; epitope; vaccine; dendritic cell;

KW tumour infiltrating lymphocyte; TIL; human leukocyte antigen; HLA;

KW cytostatic; human.

XX

OS Homo sapiens.

XX

PN WO200264057-A2.

XX

PD 22-AUG-2002.

XX

PF 15-FEB-2002; 2002WO-US05212.

XX

PR 15-FEB-2001; 2001US-268687P.

XX

XX (BAYU) BAYLOR COLLEGE MEDICINE.

PA

XX Wang R;

XX

DR WPI; 2002-627577/67.

XX

PT Novel composition for treating a disease in an animal, comprises an

PT immune effector cell and cell penetrating peptide associated with an

PT antigen or antibody -

XX

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:03:19 ; Search time 81 Seconds  
(without alignments)  
21.555 Million cell updates/sec

Title: US-09-807-512-11  
Perfect score: 50  
Sequence: 1 MLMAQELAFLL 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 271375

Minimum DB seq length: 0  
Maximum DB seq length: 11

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*  
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	100.0	11	21	CTL epitope-1 of h
2	50	100.0	11	23	Tumour antigen LAG
3	46	92.0	10	23	Human CAMEL class
4	45	90.0	10	21	CTL epitope-2 of h
5	27	54.0	9	23	Human tumour anti
6	24	48.0	9	13	Novel peptide GGF-
7	24	48.0	9	15	Cell proliferation
8	24	48.0	9	15	30-36 kD glial gro
9	24	48.0	9	17	GGF I tryptic pept

10	24	48.0	10	16	AA67176	Bovine glial cell
11	24	48.0	10	17	AAW03313	Tyrosinase peptide
12	24	48.0	10	17	AA866642	GGF I tryptic pept
13	24	48.0	10	23	AAU71255	Human MHC class I
14	24	48.0	11	22	AAW00199	Human protein frag
15	24	48.0	11	22	AAW00200	Human protein frag
16	23	46.0	9	23	ABB96933	Human tumour anti
17	23	46.0	10	23	ABB96936	Human tumour anti
18	23	46.0	11	23	AAE25495	CAPT-22 tryptic pe
19	22	44.0	8	23	ABB79903	Human BM40 basemen
20	22	44.0	9	20	AAW84253	Antigenic peptide
21	22	44.0	9	20	AAW84219	Antigenic peptide
22	22	44.0	9	20	AAW84220	Antigenic peptide
23	22	44.0	9	21	AAV70857	CAMEL10 immunogen
24	22	44.0	9	21	AAV70859	CAMEL17 immunogen
25	22	44.0	9	21	AAV66341	HLA-A24-binding HI
26	22	44.0	9	22	AAG99440	PRAME derived HLA-
27	22	44.0	9	22	AAG99483	PRAME derived HLA-
28	22	44.0	9	23	AAU71531	Human MHC class I
29	22	44.0	9	23	AAU71606	Human MHC class I
30	22	44.0	9	23	AAU71739	Human MHC molecule
31	22	44.0	10	21	AAV70858	CAMEL16 immunogen
32	22	44.0	10	22	AAG99456	PRAME derived HLA-
33	22	44.0	10	23	AAU71147	Human MHC class I
34	22	44.0	10	23	AAU71358	Human MHC class I
35	22	44.0	11	20	AAV05983	Human cancer anti
36	22	44.0	11	23	AAV67409	Human ADPI tryptic
37	22	44.0	11	24	ABR00898	Bioactive synthe
38	21	42.0	9	17	AAW01006	19K apoptosis-regu
39	21	42.0	9	20	AAW94331	19K binding intera
40	21	42.0	9	20	AAW84221	Antigenic peptide
41	21	42.0	9	21	AAV09857	Immunomodulatory p
42	21	42.0	9	21	AAV09857	Immunomodulatory p
43	21	42.0	9	23	AAE20694	Human MHC class I
44	21	42.0	9	23	AAU71132	Human MHC class I
45	21	42.0	9	23	AAU71502	Human MHC molecule

## ALIGNMENTS

## RESULT 1

AAV70855  
ID AAV70855 standard; peptide; 11 AA.

XX AAV70855;

AC AAV70855;

XX 31-JUL-2000 (first entry)

DT CTL epitope-1 of human CAMEL protein.

DE CAMEL: CTL-recognised Antigen on MELanoma; cytotoxic T lymphocyte; CTL;

KW tumour-associated antigen; LAGE-1; NY-ESO-1; anticancer; melanoma;

KW human; cancer; immunotherapy; immunogenic peptide; immune response.

XX Homo sapiens.

XX WO200023584-A1.

PD 27-APR-2000.

XX 15-OCT-1999; 99WO-EP07832.

XX 16-OCT-1998; 98EP-0119583.

XX (BOEH ) BOEHRINGER INGELHEIM INT GMBH.

XX (UYHO-) UNIV HOSPITAL LEIDEN.

XX Schrier PI, Aarnoudse CA, Helder K, Klade C;

XX WPI; 2000-339685/29.

XX Tumor-associated antigen useful for cancer immunotherapy is encoded by

**THIS PAGE BLANK (USPTO)**

Tue Oct 7 15:52:33 2003

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 3700  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1400  
; LENGTH: 96  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US03-26780-1400

Query Match 10.2%; Score 59; DB 1; Length 96;  
Best Local Similarity 31.4%; Pred. No. 1.6e+02;  
Matches 22; Conservative 4; Mismatches 40; Indels 4; Gaps 2;

QY 27 RAAEVPGAQGGQGRGEEAP---RGVRMAYPLLRRMEGAPAGPGGRTAACFSCTSRCLS 83  
Db 18 RTTAPCTSAASAPQTETRAPTRSPTRCCRPRTTCPPSPSPSTRTTATCSLSSR-WT 76

QY 84 RRPWKRWSA 93  
Db 77 TRPCRRSSA 86

RESULT 15

US-10-425-114A-55533  
; Sequence 55533, Application US/10425114A  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114A  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 55533  
; LENGTH: 101  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB143-040-F6\_FLI.pep  
US-10-425-114A-55533

Query Match 10.2%; Score 59; DB 6; Length 101;  
Best Local Similarity 30.8%; Pred. No. 1.7e+02;  
Matches 24; Conservative 9; Mismatches 37; Indels 8; Gaps 3;

QY 27 RAAEVPGAQGGQGRGEEAPRGVRMAYPLLRR---MEGAPAGPGRT---AACFSCTS 79  
Db 14 QAAEPAGICGRSGRRRRRWPSSCCSSSLRRRRRRRRACAGARASRGACPTTTA 73

QY 80 -RCLSRPWKRWSAGSC 96  
Db 74 RRCACRATAATATASC 91

Search completed: October 7, 2003, 14:04:15  
Job time : 12 secs



## RESULT 11

US-10-425-114A-46180  
; Sequence 46180, Application US/10425114A  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114A  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 46180  
; LENGTH: 86  
; TYPE: PRT  
; ORGANISM: zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 701166916\_FLI.pep  
US-10-425-114A-46180

Query Match 10.3%; Score 59.5; DB 6; Length 86;  
Best Local Similarity 47.6%; Pred. No. 1.3e+02;  
Matches 20; Conservative 3; Mismatches 14; Indels 5; Gaps 3;

QY 31 VPGAQGG--QGPRGEEAPR--GVRMAVPLLRMEGAPAGPG 68

Db 42 VAGADADLGQGGPGREPDVODGGLRGGV-ALRRPPVAAAGPG 82

## RESULT 12

PCT-US03-11231-6  
; Sequence 6, Application PC/TUS0311231  
; GENERAL INFORMATION:  
; APPLICANT: Corixa Corporation  
; APPLICANT: Day, Craig H.  
; APPLICANT: Hosken, Nancy A.  
; APPLICANT: Parsons, Joseph M.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND  
; TITLE OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION  
; FILE REFERENCE: 210121.53801PC  
; CURRENT APPLICATION NUMBER: PCT/US03/11231  
; CURRENT FILING DATE: 2003-04-09  
; NUMBER OF SEQ ID NOS: 267  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 70  
; TYPE: PRT  
; ORGANISM: Herpes simplex virus  
PCT-US03-11231-6

Query Match 10.2%; Score 59; DB 1; Length 70;  
Best Local Similarity 31.7%; Pred. No. 1.2e+02;  
Matches 19; Conservative 6; Mismatches 33; Indels 2; Gaps 1;

QY 27 RAAEVPGAQGGQPGREAP--RGVRMAVPLLRMEGAPAGPGGRTAACFCSTSLR 84

Db 1 RRARVQTTTSRRLLPRGLSRGPPRRGRPLPIRHLARRALAAAGGGGSGRPPVDCVLRDVR 60

## RESULT 13

PCT-US03-26780-1397  
; Sequence 1397, Application PC/TUS0326780  
; GENERAL INFORMATION:  
; APPLICANT: FIVEPRIME THERAPEUTICS, INC.  
; TITLE OF INVENTION: HUMAN POLYPEPTIDES ENCODED BY POLYNUCLEOTIDES AND METHODS OF  
; TITLE OF INVENTION: THEIR USE  
; FILE REFERENCE: 08940.0014-00304  
; CURRENT APPLICATION NUMBER: PCT/US03/26780

; CURRENT FILING DATE: 2003-08-28  
; PRIOR APPLICATION NUMBER: 60/406,616  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,579  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,655  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,642  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,640  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,588  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,576  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,646  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,666  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,653  
; PRIOR FILING DATE: 2002-08-29  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 3700  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1397  
; LENGTH: 96  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US03-26780-1397

Query Match 10.2%; Score 59; DB 1; Length 96;  
Best Local Similarity 31.4%; Pred. No. 1.6e+02;  
Matches 22; Conservative 4; Mismatches 40; Indels 4; Gaps 2;

QY 27 RAAEVPGAQGGQPGREAP--RGVRMAVPLLRMEGAPAGPGGRTAACFCSTSLR 83

Db 18 RTTAPPCTSAASAPQTETRAPTRSPTRCCRPRTCTCPSPWSPSTTTATCSLSSR-WT 76

QY 84 RRPWKRSWA 93

Db 77 TRPCRRSSA 86

## RESULT 14

PCT-US03-26780-1400  
; Sequence 1400, Application PC/TUS0326780  
; GENERAL INFORMATION:  
; APPLICANT: FIVEPRIME THERAPEUTICS, INC.  
; TITLE OF INVENTION: HUMAN POLYPEPTIDES ENCODED BY POLYNUCLEOTIDES AND METHODS OF  
; TITLE OF INVENTION: THEIR USE  
; FILE REFERENCE: 08940.0014-00304  
; CURRENT APPLICATION NUMBER: PCT/US03/26780  
; CURRENT FILING DATE: 2003-08-28  
; NUMBER OF SEQ ID NOS: 60/406,616  
; PRIOR APPLICATION NUMBER: 60/406,616  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,579  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,655  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,642  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,640  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,588  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,576  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,646  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,666  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/406,653  
; PRIOR FILING DATE: 2002-08-29

Db 45 RAAPRAPTAGARRRRRRPWPAAASRTDGPGRSRPGRGGRETA 94

RESULT 7

PCT-US03-24669-324  
; Sequence 324, Application PC/TUS0324669  
; GENERAL INFORMATION:  
; APPLICANT: diadexus, Inc.  
; APPLICANT: Macina, Roberto  
; APPLICANT: Salceda, Susana  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Sun, Yongming  
; APPLICANT: Turner, Leah  
; TITLE OF INVENTION: Compositions and Methods Relating to Ovarian Specific Genes and  
; FILE REFERENCE: DEX-0413  
; CURRENT APPLICATION NUMBER: PCT/US03/24669  
; CURRENT FILING DATE: 2003-08-06  
; PRIOR APPLICATION NUMBER: US 60/401,469  
; PRIOR FILING DATE: 2002-08-06  
; NUMBER OF SEQ ID NOS: 396  
; SOFTWARE: Patent version 3.1  
; SEQ ID NO 324  
; LENGTH: 108  
; TYPE: PRT  
; ORGANISM: Homo sapien  
PCT-US03-24669-324

Query Match 10.8%; Score 62.5; DB 1; Length 108;  
Best Local Similarity 27.0%; Pred. No. 96;  
Matches 31; Conservative 7; Mismatches 30; Indels 47; Gaps 6;

QY 18 LAAQRRVRAAEVPAQGOQGRGEEAPRGVMAVPLLR-----60  
Db 1 LRLSLEEA--AGEGTGRG-EGPWGRREP-----VRRPWIGGLIIVWGAING 48  
QY 61 -----EGAPAGPGGRTAACFSCTSRCLSRRPWKRSW-AGSCFPMPLSPDQG 107  
Db 49 GGSIDGGLMDWGSSD-----GGGAWGRAWGAGACCLQSALSPSSG 93

RESULT 8

US-10-425-114A-46718  
; Sequence 46718, Application US/10425114A  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114A  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 46718  
; LENGTH: 98  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700624129\_FLI.pep  
US-10-425-114A-46718

Query Match 10.7%; Score 62; DB 6; Length 98;  
Best Local Similarity 32.1%; Pred. No. 95;  
Matches 25; Conservative 8; Mismatches 37; Indels 37; Gaps 3;

QY 27 RAAEVPAQGOQGRGEEAPRGVMAVPLLR-----MEGAPAGPGRT---AACFSCTS 79  
Db 11 QAAEPAGICGSGRWRRRPSSCCSSSLRRRRRRRRACAGARAGSACPTTTA 70  
QY 80 -RCLSRRPWKRSWSAGSC 96

Db 71 RRCACRRATAAATATASC 88

RESULT 9

US-10-425-114A-66016  
; Sequence 66016, Application US/10425114A  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114A  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 66016  
; LENGTH: 98  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3957-026-E7\_FLI.pep  
US-10-425-114A-66016

Query Match 10.7%; Score 62; DB 6; Length 98;  
Best Local Similarity 32.1%; Pred. No. 95;  
Matches 25; Conservative 8; Mismatches 37; Indels 8; Gaps 3;

QY 27 RAAEVPAQGOQGRGEEAPRGVMAVPLLR-----MEGAPAGPGRT---AACFSCTS 79  
Db 11 QAAEPAGICGSGRWRRRPSSCCSSSLRRRRRRRRACAGARAGSACPTTTA 70  
QY 80 -RCLSRRPWKRSWSAGSC 96  
Db 71 RRCACRRATAAATATASC 88

RESULT 10

US-10-425-114A-51319  
; Sequence 51319, Application US/10425114A  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114A  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 51319  
; LENGTH: 109  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700218535\_FLI.pep  
US-10-425-114A-51319

Query Match 10.5%; Score 61; DB 6; Length 109;  
Best Local Similarity 40.5%; Pred. No. 1.3e+02;  
Matches 15; Conservative 4; Mismatches 10; Indels 8; Gaps 2;

QY 69 GRTAACFSCTSRCLSRRPWKRSWSAGS-----CPGMP 100  
Db 2 GSWASCWTSTSSALARRP---SWPEASSRSRPTCTCP 35

```
; SEQ ID NO 1424
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: NYSolb segment 2
US-10-296-734-1424

Query Match      26.3%; Score 152; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 OGAMLAQERRVPRAAEVPGACQGGPRGR 43
Db 1 OGAMLAQERRVPRAAEVPGACQGGPRGR 30

RESULT 3
US-10-296-734-1426
; Sequence 1426, Application US/10296734
; GENERAL INFORMATION:
; APPLICANT: Thompson, Scott A
; APPLICANT: Ramshaw, Ian A
; TITLE OF INVENTION: Synthetic molecules and uses therefor
; FILE REFERENCE: Savine
; CURRENT APPLICATION NUMBER: US/10/296,734
; PRIOR FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: AU PQ7761/00
; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1426
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: NYSolb segment 3
US-10-296-734-1426

Query Match      23.0%; Score 133; DB 6; Length 30;
Best Local Similarity 79.4%; Pred. No. 6.7e-05;
Matches 27; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 29 AEVPGAGQGPGRGREAPRGVMAVPLLRMEG 62
Db 1 AEVPGAGQGPGRGREAPRGVMAA-----RLQG 30

RESULT 4
US-10-296-734-1422
; Sequence 1422, Application US/10296734
; GENERAL INFORMATION:
; APPLICANT: Thompson, Scott A
; APPLICANT: Ramshaw, Ian A
; TITLE OF INVENTION: Synthetic molecules and uses therefor
; FILE REFERENCE: Savine
; CURRENT APPLICATION NUMBER: US/10/296,734
; PRIOR FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: AU PQ7761/00
; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1422
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: NYSolb segment 1
US-10-296-734-1422

Query Match      22.6%; Score 131; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 9.6e-05;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLMAQEALAFMLAQGAMLAQAQERRVPRA 28
Db 3 MLMAQEALAFMLAQGAMLAQAQERRVPRA 30

RESULT 5
US-10-425-114A-48446
; Sequence 48446, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 48446
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700041743_FLI.pep
US-10-425-114A-48446

Query Match      11.8%; Score 68.5; DB 6; Length 106;
Best Local Similarity 31.6%; Pred. No. 31;
Matches 30; Conservative 8; Mismatches 30; Indels 27; Gaps 5;

QY 8 LAFMLAQGAMLAQAQERRVPRAAEVPGAQ-CQGGPRGEEAPR-----GYRMAV 54
Db 20 LPWLQEEGPE-GAQERR--RRAERDGGRLAEGDGGHRRGRRHAHPVQVROEGRALAV 76
QY 55 PLLRRMEGAPAGPGGRTAACFCTSRCLSRRPWKR 89
Db 77 PPAQEERARAGGPAADSAC-----RWRR 100

RESULT 6
US-10-425-114A-50649
; Sequence 50649, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 50649
; LENGTH: 105
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700166066_FLI.pep
US-10-425-114A-50649

Query Match      11.6%; Score 67; DB 6; Length 105;
Best Local Similarity 36.0%; Pred. No. 41;
Matches 18; Conservative 2; Mismatches 30; Indels 0; Gaps 0;

QY 23 RRYPRAAEVPGAGQGPGRGREAPRGVMAVPLLRMEGAPAGPGRTA 72
```

Result No.	Query %			DB	ID	Description
	Score	Match	Length			
1	264	45.6	58	6	US-10-296-734-833	Sequence 833, App
2	192	23.0	30	6	US-10-296-734-1424	Sequence 1424, Ap
3	133	23.0	30	6	US-10-296-734-1426	Sequence 1426, Ap
4	131	22.6	30	6	US-10-296-734-1422	Sequence 1422, Ap
5	68.5	11.6	106	6	US-10-425-114A-48446	Sequence 48446, A
6	67	11.6	105	6	US-10-425-114A-50649	Sequence 50649, A
7	62.5	10.8	108	1	PCT-US03-24669-324	Sequence 324, App
8	62	10.7	98	6	US-10-425-114A-46718	Sequence 46718, A
9	62	10.7	98	6	US-10-425-114A-66016	Sequence 66016, A
10	61	10.5	109	6	US-10-425-114A-51319	Sequence 51319, A
11	59.5	10.3	86	6	US-10-425-114A-46180	Sequence 46180, A
12	59	10.2	70	1	PCT-US03-11231-6	Sequence 6, Appli
13	59	10.2	96	1	PCT-US03-26780-1397	Sequence 1397, Ap
14	59	10.2	96	1	PCT-US03-26780-1400	Sequence 1400, Ap
15	59	10.2	96	1	PCT-US03-26780-1397	Sequence 1397, Ap
16	59	10.2	102	6	US-10-425-114A-45533	Sequence 45533, A
17	59	10.2	102	6	US-10-425-114A-45162	Sequence 45162, A
18	59	10.2	104	6	US-10-425-114A-42534	Sequence 42534, A
19	58.5	10.1	109	6	US-10-425-114A-42534	Sequence 42534, A
20	58.5	10.1	96	5	US-09-638-693-192	Sequence 192, App
21	58	10.0	78	5	US-09-873-22A-192	Sequence 192, App
22	57	9.8	107	6	PCT-US02-24483-204	Sequence 204, App
23	57	9.8	107	6	US-10-296-734-1428	Sequence 1428, Ap
24	55.5	9.6	46	1	PCT-US02-31375-53	Sequence 53, Appl
25	55.5	9.6	46	1	PCT-US02-31375-55	Sequence 55, Appl
26	55.5	9.6	106	5	US-09-979-932A-1249	Sequence 1249, App

**THIS PAGE BLANK (USPTO)**

;  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 28520  
; LENGTH: 88  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-10-419-128-28520

Query Match 13.8%; Score 80; DB 30; Length 88;  
Best Local Similarity 45.9%; Pred. No. 38;  
Matches 28; Conservative 4; Mismatches 19; Indels 10; Gaps 4;

QY 42 GREEAPRGVR-MAVPLRRMEGAP--AGP-----GRTAACFCSTSRC--LSRRPWKRSW 91  
||| ||| | : ||| | ||| ||| : ||| | ||| : ||| : |||  
Db 23 GRSAPGWRGRSSPNARRTGPAAAGPPCVAGSGSACRSRSRGRRRVPRRST 82

QY 92 S 92  
|  
Db 83 S 83

RESULT 15  
US-10-437-963-119376  
; Sequence 119376, Application US/10437963  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 119376  
; LENGTH: 84  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(84)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_2259C.1.pep  
US-10-437-963-119376

Query Match 13.5%; Score 78; DB 30; Length 84;  
Best Local Similarity 40.3%; Pred. No. 55;  
Matches 27; Conservative 1; Mismatches 27; Indels 12; Gaps 3;

QY 33 GAQGGGPRGEEAPRGVRMAYVPLLRMEGAPAGPGGRTAACFCSTSRCLSR--PWK 88  
||| ||| | ||| | ||| ||| : ||| | ||| : ||| : |||  
Db 11 GGVGAGSR-RSXAGRAAAA-----AGAGGGGRTAAARSCSRSSAARGAPTWR 62

QY 89 RWSAGS 95  
|  
Db 63 RRRRAGS 69

Search completed: October 7, 2003, 14:11:01  
Job time : 400 secs



```
QY 1 MLMAQEALAFMLAQGAMLAQAQRRVPRAAEVPGAGGQGGPRGEEAPRGVMAVPLLRM 60
    |||||
Db 1 MLMAQEALAFMLAQGAMLAQAQRRVPRAAEVPGAGGQGGPRGEEAPRGVMAA-----RL 56

QY 61 EG 62
    :|
Db 57 QG 58

RESULT 6
US-09-807-512-10
; Sequence 10, Application US/09807512
; GENERAL INFORMATION:
; APPLICANT: Schrier, Peter I.
; APPLICANT: Aarnoudse, Corlien
; APPLICANT: Heider, Karl-Heinz
; APPLICANT: Klade, Christoph
; TITLE OF INVENTION: Camel, An Alternative Translation Product of the Tumor
; FILE REFERENCE: 0652.2200000
; CURRENT APPLICATION NUMBER: US/09/807,512
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: PCT/EP99/07832
; PRIOR FILING DATE: 1998-10-15
; PRIOR APPLICATION NUMBER: EP 98119583.7
; PRIOR FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-807-512-10

Query Match 45.6%; Score 264; DB 23; Length 58;
Best Local Similarity 88.7%; Pred. No. 1.1e-15;
Matches 55; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

QY 1 MLMAQEALAFMLAQGAMLAQAQRRVPRAAEVPGAGGQGGPRGEEAPRGVMAVPLLRM 60
    |||||
Db 1 MLMAQEALAFMLAQGAMLAQAQRRVPRAAEVPGAGGQGGPRGEEAPRGVMAA-----RL 56

QY 61 EG 62
    :|
Db 57 QG 58

RESULT 7
US-09-529-206-56
; Sequence 56, Application US/09529206
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America,as
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; FILE REFERENCE: 20264269PC
; CURRENT APPLICATION NUMBER: US/09/529,206
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 56
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206-56

Query Match 19.9%; Score 115; DB 19; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.0083;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLMAQEALAFMLAQGAMLAQAQRRV 25
    |||||
```

```
Db 1 MLMAQEALAFMLAQGAMLAQAQRRV 25

RESULT 8
US-09-529-206A-56
; Sequence 56, Application US/09529206A
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; FILE REFERENCE: 20264269US1
; CURRENT APPLICATION NUMBER: US/09/529,206A
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: PCT/US98/19609
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206A-56

Query Match 19.9%; Score 115; DB 19; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.0083;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLMAQEALAFMLAQGAMLAQAQRRV 25
    |||||
Db 1 MLMAQEALAFMLAQGAMLAQAQRRV 25

RESULT 9
US-09-529-206B-56
; Sequence 56, Application US/09529206B
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; FILE REFERENCE: 20264269US1
; CURRENT APPLICATION NUMBER: US/09/529,206B
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: PCT/US98/19609
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206B-56

Query Match 19.9%; Score 115; DB 19; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.0083;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLMAQEALAFMLAQGAMLAQAQRRV 25
    |||||
Db 1 MLMAQEALAFMLAQGAMLAQAQRRV 25

RESULT 10
US-09-529-206D-56
; Sequence 56, Application US/09529206D
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; APPLICANT: Rosenberg, Steven
```



Query Match	100.0%;	Score 579;	DB 23;	Length 109;
Best Local Similarity	100.0%;	Pred. No. 2e-43;		
Matches 109;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 1 MLMAQEALAFMLAQGAMLAAQERRVPRAAEVPGAQQGQPRGEEAPRGVMAVPLLRM 60

db

1 MLMAQEALAF~~LM~~AAQGA~~ML~~AAQERRVPRAAEVPGAQQGQPRGEEAPRGV~~RM~~AVPL~~LR~~RM 60

QY 61 EGAPAGPGGRTAACFSTSRCLSRPWWKRSWSAGSCPGMHLSPDQGRF 109

Db 61 EGAPAGGGRTAACFCTSRCLRRPWKRSWSAGSCPGMPLHSPDQGRF 109

## RESULT 2

```

US-09-529-206-5
; Sequence 5, Application US/09529206
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America as
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; TITLE OF INVENTION: Encoding Same
; FILE REFERENCE: 20264269PC
; CURRENT APPLICATION NUMBER: US/09/529,206
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206-5

```

Query Match 45.6%; Score 264; DB 19; Length 58;  
Best Local Similarity 88.7%; Pred. No. 1.1e-15;  
Matches 55; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

QY 1 MLMAQEALAFMLMAQGAMLAQERRVPRAAEVPGAQQGQPRGEEAPRGVRMAVPLLRM 60

db 1 MLMAQEALAFMLMAQGAMLAQERRVPRAAEVPGAQGGQGRGEEAPRGVMAA ---RL 56

QY 61 EG 62

Db 57 QG 58

### RESULTS

```

US-09-529-206A-5
; Sequence 5, Application US/09529206A
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; TITLE OF INVENTION: Encoding Same
; FILE REFERENCE: 20264269US1
; CURRENT APPLICATION NUMBER: US/09/529, 206A
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: PCT/US98/19609
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206A-5

```

Query Match 45.6%; Score 264; DB 19; Length 58;  
Best Local Similarity 88.7%; Pred. No. 1.1e-15;  
Matches 55; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

Qy 1 MLMQEALFLMQAGMLAAQERRVPRAAEVPGAQGGQGRGEEAPRGVMAVPLLRM 60  
|||||  
Db 1 MLMQEALFLMQAGMLAAQERRVPRAAEVPGAQGGQGRGEEAPRGVMAA---RL 56

QY 61 EG 62

Db 57 QG 58

## RESULT 4

```

US-09-529-206B-5
; Sequence 5, Application US/09529206B
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; TITLE OF INVENTION: Encoding Same
; FILE REFERENCE: 20264269US1
; CURRENT APPLICATION NUMBER: US/09/529,206B
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: PCT/US98/19609
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206B-5

```

Query Match 45.6%; Score 264; DB 19; Length 58;  
Best Local Similarity 88.7%; Pred. No. 1.1e-15;  
Matches 55; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

1 MLMAQEALAF~~LM~~AQGA~~ML~~AAQERRVPRAAEVPGAQQGQPRGEEAPRGVRMAVPLLRM 60

Db 1 MLMAQEALFLMAQGAMLAQERRVPRAAEVPGAQQGPRGEEAPRGVRMAA ---RL 56

QY 61 EG 62

Db 57 QG 58

## RESULT 5

```

US-09-529-206D-5
; Sequence 5, Application US/09529206D
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; TITLE OF INVENTION: Encoding Same
; FILE REFERENCE: 218791
; CURRENT APPLICATION NUMBER: US/09/529,206D
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: PCT/US98/19609
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206D-5

```

Query Match	45.6%	Score 264;	DB 19;	Length 58;
Best Local Similarity	88.7%	Pred. No. 1.le-15;		
Matches 55;	Conservative	2;	Mismatches 1;	Indels 4;
				Gaps 1;

Best local similarity 66.7%; Rec. NO. 1.1E 15;  
Matches 55; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:00:59 ; Search time 398 Seconds  
(without alignments)  
249.199 Million cell updates/sec

Title: US-09-807-512-2  
Perfect score: 579  
Sequence: 1 MIMQALAFMAQAMLA.....SWSAGSCPMHLSPDQGRF 109

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues 3647585

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 109

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents AA\_Main:\*

1: /cgn2\_6/ptodata/1/paa/PCTUS\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US05\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US080\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep.\*  
8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep.\*  
9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep.\*  
10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep.\*  
11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep.\*  
12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep.\*  
13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep.\*  
14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep.\*  
15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep.\*  
16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep.\*  
17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep.\*  
18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep.\*  
19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep.\*  
20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep.\*  
21: /cgn2\_6/ptodata/1/paa/US097A\_COMB.pep.\*  
22: /cgn2\_6/ptodata/1/paa/US097B\_COMB.pep.\*  
23: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep.\*  
24: /cgn2\_6/ptodata/1/paa/US099A\_COMB.pep.\*  
25: /cgn2\_6/ptodata/1/paa/US099B\_COMB.pep.\*  
26: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep.\*  
27: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep.\*  
28: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep.\*  
29: /cgn2\_6/ptodata/1/paa/US103\_COMB.pep.\*  
30: /cgn2\_6/ptodata/1/paa/US104\_COMB.pep.\*  
31: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep.\*  
32: /cgn2\_6/ptodata/1/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	579	100.0	109	23	US-09-807-512-2 Sequence 2, Appli

2	264	45.6	58	19	US-09-529-206-5	Sequence 5, Appli
3	264	45.6	58	19	US-09-529-206A-5	Sequence 5, Appli
4	264	45.6	58	19	US-09-529-206B-5	Sequence 5, Appli
5	264	45.6	58	19	US-09-529-206D-5	Sequence 5, Appli
6	264	45.6	58	23	US-09-807-512-10	Sequence 10, Appli
7	115	19.9	25	19	US-09-529-206-56	Sequence 56, Appli
8	115	19.9	25	19	US-09-529-206A-56	Sequence 56, Appli
9	115	19.9	25	19	US-09-529-206B-56	Sequence 56, Appli
10	115	19.9	25	19	US-09-529-206D-56	Sequence 56, Appli
11	92.5	16.0	101	1	PCT-US01-08631-41335	Sequence 41335, A
12	82.5	14.2	98	20	US-09-620-111B-4763	Sequence 4763, Ap
13	80	13.8	88	29	US-10-366-683-28520	Sequence 28520, A
14	80	13.8	88	30	US-10-419-128-28520	Sequence 28520, A
15	78	13.5	84	30	US-10-437-963-119376	Sequence 119376, A
16	78	13.5	99	21	US-09-708-427-69694	Sequence 69694, A
17	77	13.3	80	27	US-10-104-545-38	Sequence 38, Appli
18	76	13.1	107	17	US-09-383-126-22	Sequence 22, Appli
19	75	13.0	102	30	US-10-405-027-5529	Sequence 5529, Ap
20	74.5	12.9	90	29	US-10-366-683-32052	Sequence 32052, A
21	74.5	12.9	90	30	US-10-419-128-32052	Sequence 32052, A
22	74	12.8	105	21	US-09-708-427-65713	Sequence 65713, A
23	74	12.8	108	32	US-60-453-050-10768	Sequence 10768, A
24	74	12.8	108	32	US-60-453-135-10768	Sequence 10768, A
25	74	12.8	108	32	US-60-466-412-10768	Sequence 10768, A
26	73.5	12.7	94	1	PCT-US01-02723-1098	Sequence 1098, Ap
27	73.5	12.7	99	20	US-09-617-682A-8813	Sequence 8813, Ap
28	73.5	12.7	108	20	US-09-617-682A-8812	Sequence 8812, Ap
29	72.5	12.5	81	30	US-10-437-963-195393	Sequence 195393, A
30	72.5	12.5	95	20	US-09-617-681A-6686	Sequence 6686, Ap
31	72.5	12.5	105	18	US-09-489-039A-7228	Sequence 7228, Ap
32	72.5	12.5	105	30	US-10-446-203-7228	Sequence 7228, Ap
33	72	12.4	94	19	US-09-513-996A-20581	Sequence 20581, A
34	72	12.4	105	29	US-10-366-683-32394	Sequence 32394, A
35	72	12.4	105	30	US-10-419-128-32394	Sequence 32394, A
36	71	12.3	87	23	US-09-891-983A-8	Sequence 8, Appli
37	71	12.3	87	27	US-10-173-480-8	Sequence 8, Appli
38	70.5	12.2	104	1	PCT-US01-08631-54953	Sequence 54953, A
39	70.5	12.2	106	1	PCT-US01-08631-57662	Sequence 57662, A
40	70	12.1	73	1	PCT-US02-32727-9671	Sequence 9671, Ap
41	70	12.1	73	25	US-09-978-825-9671	Sequence 9671, Ap
42	70	12.1	73	26	US-10-057-498-9671	Sequence 9671, Ap
43	70	12.1	84	1	PCT-US01-08631-33076	Sequence 33076, A
44	70	12.1	93	23	US-09-834-366-26339	Sequence 26339, A
45	70	12.1	93	32	US-60-197-873-26339	Sequence 26339, A

ALIGNMENTS

RESULT 1

US-09-807-512-2  
Sequence 2, Application US/09807512  
GENERAL INFORMATION:  
APPLICANT: Schrier, Peter I.  
APPLICANT: Aarnoudse, Corlien  
APPLICANT: Heider, Karl-Heinz  
APPLICANT: Klade, Christoph  
TITLE OF INVENTION: Camel, An Alternative Translation Product of the Tumor  
TITLE OF INVENTION: Antigen-Lage 1  
FILE REFERENCE: 0652, 2200000  
CURRENT APPLICATION NUMBER: US/09/807,512  
CURRENT FILING DATE: 2001-04-16  
PRIOR APPLICATION NUMBER: PCT/EP99/07832  
PRIOR FILING DATE: 1999-10-15  
PRIOR APPLICATION NUMBER: EP 98119583.7  
PRIOR FILING DATE: 1998-10-16  
NUMBER OF SEQ ID NOS: 28  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 109  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-807-512-2

**THIS PAGE BLANK (USPTO)**

US-09-867-550-970

Query Match 10.4%; Score 60.5; DB 9; Length 76;  
Best Local Similarity 34.3%; Pred. No. 1.2e+02;  
Matches 24; Conservative 3; Mismatches 18; Indels 25; Gaps 5;  
QY 15 GAMLAAQERYPRAE-----VPGAQGOQPRGREAPGRVMAVPLLRMRMGAPAGP 67  
: ||| | | || | ||| : ||| | | ||| : ||| | | |||  
Db 15 GSALAGQ-ARAPRGAHSGTRLRPTGAE-----ARRERPRQRPHAP-----PFGGFSGP 63  
QY 68 -----GGR 70  
|||  
Db 64 RHHQGAGGR 73

RESULT 15  
US-09-220-920-50  
; Sequence 50, Application US/09220920  
; Patent No. US20020002269A1  
; GENERAL INFORMATION:  
; APPLICANT: Milbrandt, Jeffrey D.  
; APPLICANT: Baloh, Robert H.  
; TITLE OF INVENTION: Artemin, A No. US20020002269A1el Neurotrophic Factor  
; FILE REFERENCE: 6029-7996  
; CURRENT APPLICATION NUMBER: US/09/220,920  
; CURRENT FILING DATE: 1998-12-24  
; EARLIER APPLICATION NUMBER: 09/163,283  
; EARLIER FILING DATE: 1998-09-29  
; EARLIER APPLICATION NUMBER: 60/108,148  
; EARLIER FILING DATE: 1998-11-12  
; EARLIER APPLICATION NUMBER: 09/218,698  
; EARLIER FILING DATE: 1998-12-22  
; NUMBER OF SEQ ID NOS: 120  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 50  
; LENGTH: 68  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-220-920-50

Query Match 10.4%; Score 60; DB 9; Length 68;  
Best Local Similarity 35.6%; Pred. No. 1.2e+02;  
Matches 26; Conservative 5; Mismatches 26; Indels 16; Gaps 5;  
QY 38 QGPRGREAPGRVMAVPLLRMEGAPAG--PGGRTACFCSTSRCLSRPWRKWSAGS 95  
: | | | | : | : ||| ||||| : | | | | : |  
Db 7 RSPAPREGPP-----PVL-----ASPAGHLPFGRTARW--CSGRA-RRPPQPSRPAPP 52  
QY 96 CPGMPLSPDQGR 108  
| | | | |||  
Db 53 PPAPPSALPRGR 65

Search completed: October 7, 2003, 14:12:14  
Job time : 66 secs

```
Db 24 RERRGAWRTPGNDGSMRERKEAPWDAGRVLGFLQRTGATAGSPDRPRAGGQA 83
Qy 72 AACFSTSRCLSRPWR 89
Db 84 RVRCAARPRSLTIVTPHR 101

RESULT 11
US-09-925-299-877
; Sequence 877, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 877
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (100)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-877

Query Match 10.5%; Score 61; DB 9; Length 102;
Best Local Similarity 29.3%; Pred. No. 1.5e+02;
Matches 24; Conservative 9; Mismatches 27; Indels 22; Gaps 6;

Qy 33 GAQGQGGP-----RGREAPRGRVMAVP--LLRRMEGAPAGGRTAAFCSTSRCLSR 84
Db 26 GWGPEGVPSSVLRGRATCSNGICILAPHLSPAESFSPK--KSCHCFFL----- 76
Qy 85 RPWKRWSAGSCPGMPLSPDQ 106
Db 77 -PGKNW---TLPG-DRLKPEQ 93

RESULT 12
US-09-925-299-877
; Sequence 877, Application US/09925299
; Patent No. US20030040617A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 877
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (100)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-877

Db 24 RERRGAWRTPGNDGSMRERKEAPWDAGRVLGFLQRTGATAGSPDRPRAGGQA 83
Qy 72 AACFSTSRCLSRPWR 89
Db 84 RVRCAARPRSLTIVTPHR 101

RESULT 13
US-10-106-698-5898
; Sequence 5898, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypept
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 5898
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-106-698-5898

Query Match 10.5%; Score 61; DB 15; Length 102;
Best Local Similarity 29.3%; Pred. No. 1.5e+02;
Matches 24; Conservative 9; Mismatches 27; Indels 22; Gaps 6;

Qy 33 GAQGQGGP-----RGREAPRGRVMAVP--LLRRMEGAPAGGRTAAFCSTSRCLSR 84
Db 26 GWGPEGVPSSVLRGRATCSNGICILAPHLSPAESFSPK--KSCHCFFL----- 76
Qy 85 RPWKRWSAGSCPGMPLSPDQ 106
Db 77 -PGKNW---TLPG-DRLKPEQ 93

RESULT 14
US-09-867-550-970
; Sequence 970, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells a
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 970
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Homo sapiens
```





Matches 16; Conservative 3; Mismatches 9; Indels 7; Gaps 2;

QY 64 PAGPGGRTAACFCTSRCLSRRPWKR-SWSAGSCP 97  
|| ||| :| | ||| |||  
Db 7 PAAANMRTAMGTC-----KRPGRGSWSGSCP 35

RESULT 6  
US-09-925-300-1090  
; Sequence 1090, Application US/09925300  
; Patent No. US20020151681A1  
; GENERAL INFORMATION:  
; APPLICANT: Craig Rosen,  
; APPLICANT: Steve Ruben  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA101  
; CURRENT APPLICATION NUMBER: US/09/925,300  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05988  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1890  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1090  
; LENGTH: 96  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (23)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (85)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-925-300-1090  
Query Match 11.1%; Score 64.5; DB 10; Length 96;  
Best Local Similarity 25.0%; Pred. No. 66;  
Matches 22; Conservative 5; Mismatches 44; Indels 17; Gaps 2;

QY 26 PRAAEVPGAQ-----GQGGPRGEEAPRGVRYMAVPLLRMEGAPAGPGGRTA 72  
| | | | | : | | | | | : | | | | |  
Db 3 PACAREPGLPYLQVPNLRXLXSLPQRTTSPPEG-----LPQLRERSRLGLPGCAPG 58

QY 73 ACFCTSRCLSRRPWKRWSAGSCPGMP 100  
| | | | | : | | | | | : | | | | |  
Db 59 AGSDVVSSPLRTGTPARSSWPPSPRAPSPX 86

RESULT 7  
US-10-156-761-13306  
; Sequence 13306, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 13306  
; LENGTH: 97

us-09-807-512-2-szlm.rapb

US-09-864-761-35691  
; Sequence 35691, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aecomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 35691  
; LENGTH: 53  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC006391.7  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.2  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.5  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.6  
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.6  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.1  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.5  
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.6  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8  
US-09-864-761-35691  
Query Match 11.1%; Score 64.5; DB 9; Length 53;  
Best Local Similarity 45.7%; Pred. No. 37;





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:02:44 ; Search time 65 Seconds  
(without alignments)  
265.311 Million cell updates/sec

Title: US-09-807-512-2  
Perfect score: 579  
Sequence: 1 MLMQAELAFMLAQAMLA.....SWSAGSCPMPLSPDQGRF 109

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues  
Total number of hits satisfying chosen parameters: 224079

Minimum DB seq length: 0  
Maximum DB seq length: 109

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09D\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	71	12.3	87	11 US-09-891-983A-8	Sequence 8, Appli
2	71	12.3	87	12 US-10-173-480-8	Sequence 8, Appli
3	69	11.9	78	9 US-09-864-761-48354	Sequence 48354, A
4	66.5	11.5	107	9 US-09-220-920-52	Sequence 52, Appl
5	64.5	11.1	53	9 US-09-864-761-35691	Sequence 35691, A
6	64.5	11.1	96	10 US-09-925-300-1090	Sequence 1090, Ap
7	64	11.1	97	15 US-10-156-761-13306	Sequence 13306, A
8	64	11.1	99	12 US-09-933-767-1133	Sequence 1133, Ap
9	64	11.1	99	15 US-10-023-282-1133	Sequence 1133, Ap
10	63.5	11.0	104	15 US-10-106-698-4905	Sequence 4905, Ap
11	61	10.5	102	9 US-09-925-299-877	Sequence 877, App
12	61	10.5	102	11 US-09-925-299-877	Sequence 877, App
13	61	10.5	102	15 US-10-106-698-5898	Sequence 5898, Ap
14	60.5	10.4	76	9 US-09-867-550-970	Sequence 970, App
15	60	10.4	68	9 US-09-220-920-50	Sequence 50, Appli

16	60	10.4	75	11	US-09-891-983A-6	Sequence 6, Appli
17	60	10.4	75	12	US-10-173-480-6	Sequence 6, Appli
18	60	10.4	104	10	US-09-764-846-153	Sequence 153, App
19	60	10.4	104	15	US-10-091-483-153	Sequence 153, App
20	60	10.4	104	15	US-10-106-698-4992	Sequence 4992, Ap
21	59.5	10.3	57	9	US-09-864-761-35843	Sequence 35843, A
22	59.5	10.3	88	12	US-09-933-767-730	Sequence 730, App
23	59.5	10.3	88	15	US-10-023-282-730	Sequence 730, App
24	59.5	10.3	93	10	US-09-731-872-462	Sequence 462, App
25	59.5	10.3	93	12	US-09-876-997-462	Sequence 462, App
26	59.5	10.3	107	9	US-09-864-761-35252	Sequence 35252, A
27	59	10.2	70	9	US-09-894-998-6	Sequence 6, Appli
28	59	10.2	70	12	US-10-200-562-6	Sequence 6, Appli
29	59	10.2	70	12	US-10-237-551-6	Sequence 6, Appli
30	59	10.2	70	15	US-10-121-988-6	Sequence 6, Appli
31	59	10.2	80	11	US-09-809-391-716	Sequence 716, App
32	59	10.2	80	12	US-09-882-171-716	Sequence 716, App
33	59	10.2	85	14	US-10-016-481-16	Sequence 16, Appl
34	59	10.2	85	15	US-10-323-157-16	Sequence 16, Appl
35	59	10.2	86	14	US-10-016-481-3	Sequence 3, Appli
36	59	10.2	86	15	US-10-016-481-17	Sequence 3, Appli
37	59	10.2	86	15	US-10-323-157-3	Sequence 17, Appl
38	59	10.2	86	15	US-10-323-157-17	Sequence 17, Appl
39	59	10.2	87	14	US-10-016-481-18	Sequence 18, Appl
40	59	10.2	87	15	US-10-323-157-18	Sequence 18, Appl
41	59	10.2	89	14	US-10-016-481-15	Sequence 15, Appl
42	59	10.2	89	15	US-10-323-157-15	Sequence 15, Appl
43	59	10.2	105	9	US-09-989-722-371	Sequence 371, App
44	59	10.2	105	9	US-09-989-723-371	Sequence 371, App
45	59	10.2	105	9	US-09-989-279-371	Sequence 371, App

ALIGNMENTS

RESULT 1  
US-09-891-983A-8  
; Sequence 8, Application US/09891983A  
; Publication No. US20030108858A1  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Shah, Dinesh O.  
; APPLICANT: Dawson, George A.  
; APPLICANT: Muerhoff, A. Scott  
; APPLICANT: Jiang, Lily  
; APPLICANT: Gutierrez, Robin A.  
; APPLICANT: Leary, Thomas P.  
; APPLICANT: Desai, Suresh  
; APPLICANT: Stewart, James L.  
; TITLE OF INVENTION: Methods For The simultaneous Detection  
; TITLE OF INVENTION: Of HCV Antigens And HCV Antibodies  
; FILE REFERENCE: 6821.US.01  
; CURRENT APPLICATION NUMBER: US/09/891,983A  
; CURRENT FILING DATE: 2001-06-26  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: FASTSEQ For Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 87  
; TYPE: PRT  
; ORGANISM: Hepatitis C Virus  
US-09-891-983A-8

Query Match 12.3%; Score 71; DB 11; Length 87;  
Best Local Similarity 37.0%; Pred. No. 15;  
Matches 20; Conservative 9; Mismatches 19; Indels 6; Gaps 3;  
QY 21 QERRVPRAAVEPQAQQGPR-GREA----PRGVMAVPLLLRRMGAP-AGPG 68  
Db 14 QDVKPPGGQIVGLLPRLGRGRLKTSERSQPRGRQIPKARRPESGRTWAQPG 67

RESULT 2  
US-10-173-480-8



QY 74 CFSCTSRCLSRPWRKRSWSAG-SCPG 98  
Db 87 -----RRP-----CSARGCPG 98

## RESULT 13

US-07-972-032-59  
; Sequence 59, Application US/07972032  
; Patent No. 5496712  
; GENERAL INFORMATION:  
; APPLICANT: Cappello, Joseph  
; APPLICANT: Ferrari, Franco A.  
; TITLE OF INVENTION: HIGH MOLECULAR WEIGHT COLLAGEN-LIKE  
; TITLE OF INVENTION: PROTEIN POLYMERS  
; NUMBER OF SEQUENCES: 85  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Bertram I. Rowland  
; STREET: 4 Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: CA 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; APPLICATION NUMBER: US/07/972,032  
; FILING DATE: 19921105  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/791,960  
; FILING DATE: 12-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rowland, Bertram I.  
; REGISTRATION NUMBER: 20,015  
; REFERENCE/DOCKET NUMBER: A-55556-1/BIR; PROP-08-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 781-1989  
; TELEFAX: (415) 398-3249  
; INFORMATION FOR SEQ ID NO: 59:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 54 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-972-032-59

Query Match 11.1%; Score 64.5; DB 1; Length 54;  
Best Local Similarity 40.0%; Pred. No. 3.4;  
Matches 16; Conservative 2; Mismatches 9; Indels 13; Gaps 2;

QY 32 PGAQGGQGGPRGR--EEAPRGVRMAVPLLRMEGAPGPG 69  
Db 18 PGAQGPAGPGSGRGDPGPGAQ-----GPAGPGG 46

## RESULT 14

US-08-642-255-71  
; Sequence 71, Application US/08642255  
; Patent No. 5773249  
; GENERAL INFORMATION:  
; APPLICANT: CAPPELLO, Joseph  
; APPLICANT: FERRARI, Franco A.  
; TITLE OF INVENTION: High Molecular Weight Collagen-Like  
; TITLE OF INVENTION: Protein Polymers  
; NUMBER OF SEQUENCES: 135  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT  
; STREET: 4 Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: California

; COUNTRY: USA  
; ZIP: 94111-4187  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/642,255  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: ROWLAND, Bertram I.  
; REGISTRATION NUMBER: 20,015  
; REFERENCE/DOCKET NUMBER: A55556-3/BIR  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 494-8700  
; TELEFAX: (415) 494-8771  
; TELEX: 910 277299 FHT UR  
; INFORMATION FOR SEQ ID NO: 71:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 54 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-642-255-71

Query Match 11.1%; Score 64.5; DB 1; Length 54;  
Best Local Similarity 40.0%; Pred. No. 3.4;  
Matches 16; Conservative 2; Mismatches 9; Indels 13; Gaps 2;

QY 32 PGAQGGQGGPRGR--EEAPRGVRMAVPLLRMEGAPGPG 69  
Db 18 PGAQGPAGPGSGRGDPGPGAQ-----GPAGPGG 46

## RESULT 15

US-09-205-258-1133  
; Sequence 1133, Application US/09205258  
; Patent No. 6525174  
; GENERAL INFORMATION:  
; APPLICANT: Young et al.  
; TITLE OF INVENTION: 207 Human Secreted Proteins  
; FILE REFERENCE: P2007P1  
; CURRENT APPLICATION NUMBER: US/09/205,258  
; CURRENT FILING DATE: 1998-12-04  
; EARLIER APPLICATION NUMBER: PCT/US98/11422  
; EARLIER FILING DATE: 1998-06-04  
; EARLIER APPLICATION NUMBER: 60/048,885  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,375  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,881  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,880  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,896  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,020  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,876  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,895  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,884  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,894  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,971  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,964  
; EARLIER FILING DATE: 1997-06-06

US-08-642-255-94

Query Match 11.4%; Score 66; DB 1; Length 60;  
Best Local Similarity 36.7%; Pred. No. 2.6;  
Matches 22; Conservative 5; Mismatches 27; Indels 6; Gaps 2;

QY 15 GAMLAAQERRVPRAAEVPGAGQGQPGRGREAP--RGVRMAVPLLRMEGAPAGPGGRTA 72  
||| | : | ||| | | : | | : | : ||| | : |  
Db 1 GAGPAGSQAGPGLQAGPAGPGSQAGPGLQAGPAGSQAGPGL-----QGAPGAPGSGQA 56

RESULT 10

US-07-972-032-79

; Sequence 79, Application US/07972032

; Patent No. 5496712

; GENERAL INFORMATION:

; APPLICANT: Cappello, Joseph

; APPLICANT: Ferrari, Franco A.

; TITLE OF INVENTION: HIGH MOLECULAR WEIGHT COLLAGEN-LIKE

; TITLE OF INVENTION: PROTEIN POLYMERS

; NUMBER OF SEQUENCES: 85

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Bertram I. Rowland

; STREET: 4 Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: CA 94111

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/972,032

; FILING DATE: 19921105

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/07/791,960

; FILING DATE: 12-NOV-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Rowland, Bertram I.

; REGISTRATION NUMBER: 20,015

; REFERENCE/DOCKET NUMBER: A-55556-1/BIR:PROP-08-1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 781-1989

; TELEFAX: (415) 398-3249

; INFORMATION FOR SEQ ID NO: 79:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 54 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-07-972-032-79

Query Match 11.2%; Score 65; DB 1; Length 54;  
Best Local Similarity 34.5%; Pred. No. 3;  
Matches 20; Conservative 4; Mismatches 26; Indels 8; Gaps 2;

QY 15 GAMLAAQERRVPRAAEVPGAGQGQPGRGREAPRGVRMAVPLLRMEGAPAGPGGRTA 72  
||| | : | ||| | | : | | : | : ||| | : |  
Db 1 GAGPAGSQAGPGLQAGPAGPGSQAGPGLQAGPAGSQAGPGL-----GAP---GSGAPGAPGSGQA 50

RESULT 11

US-08-642-255-92

; Sequence 92, Application US/08642255

; Patent No. 5773249

; GENERAL INFORMATION:

; APPLICANT: Cappello, Joseph

; APPLICANT: FERRARI, Franco A.

; TITLE OF INVENTION: High Molecular Weight Collagen-Like

; TITLE OF INVENTION: Protein Polymers

; NUMBER OF SEQUENCES: 135

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON &amp; HERBERT

; STREET: 4 Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94111-4187

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/642,255

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: ROWLAND, Bertram I.

; REGISTRATION NUMBER: 20,015

; REFERENCE/DOCKET NUMBER: A55556-3/BIR

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 494-8700

; TELEFAX: (415) 494-8771

; TELEX: 910 277299 FHT UR

; INFORMATION FOR SEQ ID NO: 92:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 54 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-642-255-92

Query Match 11.2%; Score 65; DB 1; Length 54;

Best Local Similarity 34.5%; Pred. No. 3;

Matches 20; Conservative 4; Mismatches 26; Indels 8; Gaps 2;

QY 15 GAMLAAQERRVPRAAEVPGAGQGQPGRGREAPRGVRMAVPLLRMEGAPAGPGGRTA 72  
||| | : | ||| | | : | | : | : ||| | : |

Db 1 GAGPAGSQAGPGLQAGPAGPGSQAGPGLQAGPAGSQAGPGL-----GAP---GSGAPGAPGSGQA 50

RESULT 12

US-09-252-991A-31180

; Sequence 31180, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 31180

; LENGTH: 98

; TYPE: PRT

; ORGANISM: pseudomonas aeruginosa

US-09-252-991A-31180

Query Match 11.2%; Score 65; DB 4; Length 98;

Best Local Similarity 33.7%; Pred. No. 5.9;

Matches 29; Conservative 3; Mismatches 22; Indels 32; Gaps 6;

QY 26 PRA-----AEVPGAQGGQPRG-----REEAPRGVRMAVPLLRMEGAPAGPGGRTAA 73  
||| | : | ||| | | : | | : | : ||| | : |

Db 32 PRAGRGTA DSPGNHGFPGGGGLPPDRRRDRAGRSRRGS-----GSARGAPGPGRGADV - 86

```
Query Match 11.5%; Score 66.5; DB 4; Length 92;
Best Local Similarity 27.8%; Pred. No. 3.8;
Matches 25; Conservative 6; Mismatches 26; Indels 33; Gaps 4;

QY 13 AQGAML-----AAQERRVPRAAEVPGAQOGPGRGEEAPRGVMAVPLLRMEGAPAG 66
Db 22 AQGRTAPSAPTATAPRAAPASALPGRQRSSFFGRGWSRCV-----PAG 68

QY 67 PGRTAACFCTSLRSLRPWKRSWSAGSC 96
Db 69 P-----LAARCDPRP-----GDGSC 84

RESULT 7
US-09-220-528-52
; Sequence 52, Application US/09220528A
; Patent No. 6284540
; GENERAL INFORMATION:
; APPLICANT: Milbrandt, Jeffrey D.
; APPLICANT: Baloh, Robert H.
; TITLE OF INVENTION: Artemin, A No. 6284540el Neurotrophic Factor
; FILE REFERENCE: 6029-7998
; CURRENT APPLICATION NUMBER: US/09/220,528A
; CURRENT FILING DATE: 1998-12-24
; EARLIER APPLICATION NUMBER: 09/218,698
; EARLIER FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 60/108,148
; EARLIER FILING DATE: 1998-11-12
; EARLIER APPLICATION NUMBER: 09/163,283
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-220-528-52

Query Match 11.5%; Score 66.5; DB 3; Length 107;
Best Local Similarity 32.1%; Pred. No. 4.5;
Matches 34; Conservative 9; Mismatches 40; Indels 23; Gaps 6;

QY 5 QEALFLMAQGAAMLAQERRVPRAAEVPGAQOGPGRGEEAPRGVMAVPLLRMEGAP 64
Db 20 QPALMPTLAALLS-----VAEASLGSAPSPAPREGPP-----PVL-----ASP 61

QY 65 AG--PGRTPAACFCTSLRSLRPWKRSWSAGSCGPMHLSPDQGR 108
Db 62 AGHLPGGRTARW--CSGRA--RRPPQPSRPAPPPAPPSALPRGGR 104

RESULT 8
US-07-972-032-81
; Sequence 81, Application US/07972032
; Patent No. 5496712
; GENERAL INFORMATION:
; APPLICANT: Cappello, Joseph
; APPLICANT: Ferrari, Franco A.
; TITLE OF INVENTION: HIGH MOLECULAR WEIGHT COLLAGEN-LIKE
; TITLE OF INVENTION: PROTEIN POLYMERS
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bertram I. Rowland
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: CA 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/972.032
; FILING DATE: 19921105
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/791.960
; FILING DATE: 12-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I.
; REGISTRATION NUMBER: 20,015
; REFERENCE/DOCKET NUMBER: A-55556-1/BIR;PROP-08-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; INFORMATION FOR SEQ ID NO: 81:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 60 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-972-032-81

Query Match 11.4%; Score 66; DB 1; Length 60;
Best Local Similarity 36.7%; Pred. No. 2.6;
Matches 22; Conservative 5; Mismatches 27; Indels 6; Gaps 2;

QY 15 GAMLAQERRVPRAAEVPGAQOGPGRGEEAP--RGVMAVPLLRMEGAPAGPGGRTA 72
Db 1 GAPGAPGOGAPGLQGAPGAPGOGAPGLQGAPGAPGOGAPGL-----QGAPGAPGOGA 56

RESULT 9
US-08-642-255-94
; Sequence 94, Application US/08642255
; Patent No. 5773249
; GENERAL INFORMATION:
; APPLICANT: CAPPELLO, Joseph
; APPLICANT: FERRARI, Franco A.
; TITLE OF INVENTION: High Molecular Weight Collagen-Like
; TITLE OF INVENTION: Protein Polymers
; NUMBER OF SEQUENCES: 135
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/642,255
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: ROWLAND, Bertram I.
; REGISTRATION NUMBER: 20,015
; REFERENCE/DOCKET NUMBER: A55556-3/BIR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 494-8700
; TELEFAX: (415) 494-8771
; TELEX: 910 277299 FHT UR
; INFORMATION FOR SEQ ID NO: 94:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 60 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
```

; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 32052  
; LENGTH: 90  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-32052

Query Match 12.9%; Score 74.5; DB 4; Length 90;  
Best Local Similarity 38.6%; Pred. No. 0.51;  
Matches 27; Conservative 3; Mismatches 25; Indels 15; Gaps 5;  
QY 40 PRGR-----EAPRGVMAVPLLLRRMEGAPAGPGGRTA---ACFSCTSRCLSRPWPWKRWSW 91  
Db 2 PRASCRRESSAPAG-----SPGARRRSPVPAGSPCRSPWPACPAC---GCAAGRPGTGCW 55  
QY 92 -SAGSCPGMP 100  
Db 56 CSAGSTPGRP 65

## RESULT 3

US-09-252-991A-32394  
; Sequence 32394, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 32394  
; LENGTH: 105  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-32394

Query Match 12.4%; Score 72; DB 4; Length 105;  
Best Local Similarity 32.9%; Pred. No. 1.1;  
Matches 26; Conservative 5; Mismatches 26; Indels 22; Gaps 4;  
QY 42 GREEAPRGVMAVPLLLRRMEGAPAGPG-----GRTAACFS-----CTSRCLS---RR 85  
Db 21 GRRSPGSGSSAAPTTRRRRCGAPTGSRTSAGCGSRPCSAHRHRTGRCAAGCGATGCRR 80  
QY 86 PWTRWS-----AGSCPG 98  
Db 81 PVRRGSPRWDVAGSLAG 99

## RESULT 4

US-09-252-991A-23481  
; Sequence 23481, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 23481  
; LENGTH: 96

; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-23481

Query Match 11.8%; Score 68.5; DB 4; Length 96;  
Best Local Similarity 30.4%; Pred. No. 2.4;  
Matches 24; Conservative 7; Mismatches 31; Indels 17; Gaps 4;  
QY 28 AAEVPGAQGGQGRGEEAPRGV-----RMAVPLLLRRMEGAPGGGRTAACFSCTSRCLS 83  
Db 27 SAGLPGLLSRPGVPRRAAPAGSGLPRRSVVP-----TPAGPSARRPV--APTPOCAT 77  
QY 84 RRPWKR-----SWSAGSCPG 98  
Db 78 GRGWRPAPGSGSGRLPG 96

## RESULT 5

US-09-252-991A-19337  
; Sequence 19337, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 19337  
; LENGTH: 91  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-19337

Query Match 11.7%; Score 67.5; DB 4; Length 91;  
Best Local Similarity 34.4%; Pred. No. 2.9;  
Matches 31; Conservative 6; Mismatches 32; Indels 21; Gaps 6;  
QY 19 AAQERRVPAAEYVGAQGGQGRGEEAPR-----GYRMAVPLLLRRMEGAPGGGRTAA 73  
Db 3 AHRDGRTPVA--VPHRRGARAAGCTGCPDPDGTGTGPRAGVP--RRSPGAAAG----- 52  
QY 74 CFSCTSRCLSRPWPWKRWSW--SAGSC-PCMP 100  
Db 53 ---CSRSGAHRPAAGSGCPEAGSCRPA 79

## RESULT 6

US-09-252-991A-29573  
; Sequence 29573, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 29573  
; LENGTH: 92  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-29573

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:56:04 ; Search time 28 Seconds  
(without alignments)  
164.710 Million cell updates/sec

Title: US-09-807-512-2

Perfect score: 579

Sequence: 1 MLMQAEALFLMAQGAMLA.....SWSAGSCPGMPLSPDOGRF 109

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 232631

Minimum DB seq length: 0

Maximum DB seq length: 109

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents\_AA:\*

- 1: /cgn2\_6/ptodata/1/iaa/5A.COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B.COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A.COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B.COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PCRTUS.COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	80	13.8	88	4	US-09-252-991A-28520
2	74.5	12.9	90	4	US-09-252-991A-32052
3	72	12.4	105	4	US-09-252-991A-32394
4	68.5	11.8	96	4	US-09-252-991A-23481
5	67.5	11.7	91	4	US-09-252-991A-19337
6	66.5	11.5	92	4	US-09-252-991A-29573
7	66.5	11.5	107	3	US-09-220-528-52
8	66	11.4	60	1	US-07-972-032-81
9	66	11.4	60	1	US-08-642-255-94
10	65	11.2	54	1	US-07-972-032-79
11	65	11.2	54	1	US-08-642-255-92
12	65	11.2	98	4	US-09-252-991A-31180
13	64.5	11.1	54	1	US-07-972-032-59
14	64.5	11.1	54	1	US-08-642-255-71
15	64	11.1	99	4	US-09-205-252-1133
16	63	10.9	108	4	US-09-252-991A-28912
17	62	10.7	103	4	US-09-252-991A-21485
18	62	10.7	108	4	US-09-252-991A-17778
19	61.5	10.6	60	4	US-09-336-536-6
20	61.5	10.6	95	4	US-09-252-991A-16640
21	61	10.5	63	1	US-08-642-255-121
22	61	10.5	63	1	US-08-397-633A-32
23	61	10.5	65	1	US-08-642-255-125
24	61	10.5	65	1	US-08-397-633A-46
25	60.5	10.4	92	4	US-09-252-991A-24691
26	60	10.4	62	1	US-08-616-732A-24
27	60	10.4	62	3	US-09-037-742B-24

28	60	10.4	68	3	US-09-220-528-50	Sequence 50, Appl
29	60	10.4	77	4	US-09-252-991A-21642	Sequence 21642, A
30	60	10.4	84	1	US-07-972-032-68	Sequence 68, Appl
31	60	10.4	93	1	US-07-972-032-60	Sequence 60, Appl
32	60	10.4	93	1	US-07-972-032-61	Sequence 61, Appl
33	60	10.4	107	4	US-09-252-991A-30944	Sequence 30944, A
34	59.5	10.3	81	4	US-09-252-991A-24812	Sequence 24812, A
35	59.5	10.3	88	4	US-09-205-258-730	Sequence 730, App
36	59.5	10.3	89	4	US-09-252-991A-32804	Sequence 32804, A
37	59.5	10.3	109	4	US-09-252-991A-19878	Sequence 19878, A
38	59	10.2	70	4	US-09-894-998A-6	Sequence 6, Appl
39	59	10.2	80	4	US-09-149-476-716	Sequence 716, App
40	59	10.2	99	4	US-09-252-991A-28315	Sequence 28315, A
41	59	10.2	105	4	US-09-996-243-371	Sequence 371, App
42	59	10.2	105	4	US-09-712-529-5	Sequence 5, Appl
43	59	10.2	107	4	US-09-252-991A-27780	Sequence 27780, A
44	58.5	10.1	72	4	US-09-252-991A-30919	Sequence 30919, A
45	58.5	10.1	76	4	US-09-252-991A-25348	Sequence 25348, A

#### ALIGNMENTS

##### RESULT 1

US-09-252-991A-28520  
; Sequence 28520, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 28520  
; LENGTH: 88  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-28520

Query Match	13.8%	Score 80;	DB 4;	Length 88;
Best Local Similarity	45.9%	Pred. No. 0.13;		
Matches	28;	Conservative	4;	Mismatches 19; Indels 10; Gaps 4;
QY	42	GREENPRGVR-MAVPLLRMEGAP--AGP-----GGRTAACFCTSRC--LSRRPWKRSW	91	
Db	23	GRRSAPGGWGRSSPNARRRTGPPAAAGPPCVAGGSGSACRSRSGRRSRVPRRST	82	
QY	92	S	92	
Db	83	S	83	

##### RESULT 2

US-09-252-991A-32052  
; Sequence 32052, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run On: October 7, 2003, 13:55:14 ; Search time 94 Seconds  
(without alignments)  
299.231 Million cell updates/sec

Title: US-09-807-512-2

Perfect score: 579

Sequence: 1 MLMQAEALFLMAQAGMLAA.....SWSAGSCPGMPLSPDQGRF 109

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 167439

Minimum DB seq length: 0

Maximum DB seq length: 109

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	579	100.0	109	4	Q95987
2	264	45.6	58	4	Q95146
3	91	15.7	94	4	Q81278
4	78	13.5	105	11	Q9D5L2
5	77	13.3	108	10	Q8H321
6	70	12.1	100	4	Q9NY7
7	69	11.9	100	4	Q9UGS8
8	68.5	11.8	102	10	Q8S0Y5
9	68	11.7	86	4	Q8N493
10	68	11.7	95	5	Q8WT20
11	67	11.6	92	11	Q61022
12	67	11.6	99	15	Q90DC7
13	66.5	11.5	90	10	Q8S6N8
14	65	11.2	102	6	Q95KS1
15	64.5	11.1	85	5	Q9VR11
16	64.5	11.1	101	10	Q9AX16
					Q9ax16 oryza sativ

17	64.5	11.1	108	10	Q8H8C1
18	64	11.1	85	12	Q91KM1
19	63.5	11.0	27	4	Q9UMA6
20	63.5	11.0	99	11	Q63828
21	63.5	11.0	107	11	Q8C572
22	63	10.9	33	4	Q9UE40
23	62.5	10.8	82	6	O77539
24	62.5	10.8	108	17	Q9YEB6
25	62	10.7	100	11	Q9D475
26	61.5	10.6	85	12	Q91KK3
27	61.5	10.6	100	12	Q9QT53
28	61	10.5	96	5	Q17088
29	60.5	10.4	103	12	Q68346
30	60.5	10.4	73	16	Q9K437
31	60.5	10.4	84	4	Q14038
32	60.5	10.4	100	12	Q8QP85
33	60.5	10.4	100	12	Q8QT56
34	60.5	10.4	100	12	Q8QP87
35	60.5	10.4	100	12	Q8QP72
36	60.5	10.4	100	12	Q8QP71
37	60.5	10.4	100	12	Q8QP74
38	60.5	10.4	100	12	Q8QP80
39	60.5	10.4	100	12	Q8QP77
40	60.5	10.4	100	12	Q8QP84
41	60.5	10.4	100	12	Q8QP82
42	60.5	10.4	100	12	Q8QP78
43	60.5	10.4	100	12	Q8QP75
44	60.5	10.4	100	12	Q8QP81
45	60.5	10.4	100	12	Q8QP83

#### ALIGNMENTS

#### RESULT 1

Q95987 PRELIMINARY; PRT; 109 AA.

AC O95987;  
DT 01-MAY-1999 (TREMREL. 10, Created)  
DT 01-MAY-1999 (TREMREL. 10, Last sequence update)  
DT 01-MAY-1999 (TREMREL. 10, Last annotation update)  
DE CTL-recognized antigen ON melanoma (CAMEL).  
GN CAMEL.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Melanoma;

RA Aarnoudse C.A., Doel vanden P.B., Heemskerk B., Schrier P.I.;

RT "IL-2 induced melanoma-specific CTL recognize CAMEL, an unexpected

translation product of LAGE-1.";

RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ012835; CAA10197.1; -

DR EMBL; AJ012833; CAA10193.1; -

DR EMBL; AJ012834; CAA10195.1; -

SQ SEQUENCE 109 AA, 11689 MW; 189F85BC04C1F5F0 CRC64;

Query Match

Best Local Similarity 100.0%; Score 579; DB 4; Length 109;

Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLMQAEALFLMAQAGMLAAQERRVPRAAEVPGAGQCGQCGPRGEEAPRGVRAVPLLRM 60

Db 1 MLMQAEALFLMAQAGMLAAQERRVPRAAEVPGAGQCGQCGPRGEEAPRGVRAVPLLRM 60

QY 61 EGAPAGPGGRTAACFSCTSRCLSRPWRKFSWSAGSCPGMPLSPDQGRF 109

Db 61 EGAPAGPGGRTAACFSCTSRCLSRPWRKFSWSAGSCPGMPLSPDQGRF 109

#### RESULT 2

```
095146
ID O95146 PRELIMINARY; PRT; 58 AA.
AC O95146;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE LAGE-2ALT protein isoform.
GN LAGE-2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang R.-F., Johnston S.L., Zeng G., Topalian S.L.,
RA Schwartzentruber D.J., Rosenberg S.A.;
RT "A Breast and Melanoma-Shared Tumor Antigen: T Cell Responses to
RT Antigenic Peptides Translated from Different Open Reading Frames.";
PL J. Immunol. 161:3596-3606(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Lethe B.G.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=99454989; PubMed=10523621;
RA De Smet C., Lurquin C., Lethe B., Martelange V., Boon T.;
RT "DNA methylation is the primary silencing mechanism for a set of germ
RT line- and tumor-specific genes with a CpG-rich promoter.";
RL Mol. Cell. Biol. 19:7327-7335(1999).
DR EMBL; AF038567; AAD05203.1; -
DR EMBL; AJ275977; CAB76944.1; -
SQ SEQUENCE 58 AA; 6188 MW; ED12057564BC7EF2 CRC64;

Query Match 45.6%; Score 264; DB 4; Length 58;
Best Local Similarity 88.7%; Pred. No. 1.9e-18;
Matches 55; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

Qy 1 LMAQAEALFLMAQAGMLAAQERRVPRAAEVPGAGQGGPRGEEAPRGVMAVPLLRM 60
Db 1 LMAQAEALFLMAQAGMLAAQERRVPRAAEVPGAGQGGPRGEEAPRGVMAA-----RL 56

Qy 61 EG 62
:1
Db 57 QG 58

RESULT 3
Q81278 PRELIMINARY; PRT; 94 AA.
AC Q81278;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC015744; AAHL5744.1; -
KW Hypothetical protein.
SQ SEQUENCE 94 AA; 10329 MW; C95D33426364379D CRC64;

Query Match 15.7%; Score 91; DB 4; Length 94;
Best Local Similarity 33.0%; Pred. No. 0.14;
Matches 32; Conservative 4; Mismatches 51; Indels 10; Gaps 2;

Qy 11 LMAQAEALFLMAQAGMLAAQERRVPRAAEVPGAGQGGPRGEEAPRGVMAVPLLRMEGAPGGR 70
```

```
Db 1 MQTQAEALTAGMAGVATAAGAWTQPLRPVELPORTQVRAETPRLR-----PGGH 52
Qy 71 TAACFSCTSRCLSRPWRKRSWSAGSCPGMHPILSPDQG 107
Db 53 ECGRTYSPSACLSRPPWRKRSKSPMGPHQMP--SPTKG 87

RESULT 4
Q9D5L2 PRELIMINARY; PRT; 105 AA.
AC Q9D5L2;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE 4930426D05Rik protein.
GN 4930426D05Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kadosawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamliya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Tovo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK015205; BAB29745.1; -
DR MGD; MGI:1921894; 4930426D05Rik.
SQ SEQUENCE 105 AA; 10832 MW; 431FF01AD3435A22 CRC64;

Query Match 13.5%; Score 78; DB 11; Length 105;
Best Local Similarity 35.2%; Pred. No. 2.9;
Matches 31; Conservative 7; Mismatches 34; Indels 16; Gaps 5;

Qy 19 AAQERRVPRAAEVPGAGQGGPR--GEEAPRGVMAVPLLRMEGAPGGRTAACFS 76
Db 17 ASRAQRRPTAA--VPGL--NRKGPSCPCPARAPAGTTAAARLR-----GTRRTCPFR 65
Qy 77 CTSRCLSRPWRKRSWSAGSCPGMHPILSP 104
Db 66 CGSLCLARRGF---WIAGFIPTRLGSSP 90

RESULT 5
Q8H321 PRELIMINARY; PRT; 108 AA.
AC Q8H321;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE OSJNBA0077M12.27 protein.
GN OSJNBA0077M12.27.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Embryophyta; Streptophyta; Tracheophyta;
```

```
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OX Eubryotidae; Oryzae; Oryza.
RN NCBI_TaxID=39947;
RP SEQUENCE FROM N.A.
RA STRAIN=cv. Nipponbare;
RC Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 8, BAC
clone:OSJNBa007M12.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005301; BAC24969.1; -.
SQ SEQUENCE 108 AA; 10826 MW; 75A593DEF4EEC953 CRC64;

Query Match 13.3%; Score 77; DB 10; Length 108;
Best Local Similarity 29.3%; Pred. No. 3.7;
Matches 29; Conservative 7; Mismatches 43; Indels 20; Gaps 5;

QY 12 MAQGAMLAQERRVPR-----AAEVPGAQGGQGGPRGEEAPRGVMAVPLLRMEGAPAG 66
Db 1 MADGVLPARRQRRLPADGVLPATAAPGGAATAP-GRATAAAG----AAPASTAAGGGGA 56

QY 67 PGRTAACFCTSRCLSRPWPKRWSAG-----SCPGMP 100
Db 57 PRTREAA-----ARARVWNRRTAGGGVRRACSGEP 89

RESULT 6
Q9NY7 PRELIMINARY; PRT; 100 AA.
AC Q9NY7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE Low-voltage-activated calcium channel alpha13.2 subunit
DE (Fragment).
DE CACNA1H.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Mitman S., Agnew W.S.;
RT "Organization and Alternative Splicing of CACNA1H.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF223560; AAF60160.2; -.
FT NON_TER 100
SQ SEQUENCE 100 AA; 10266 MW; D00EE31D23DB5FD CRC64;

Query Match 12.1%; Score 70; DB 4; Length 100;
Best Local Similarity 30.6%; Pred. No. 16;
Matches 30; Conservative 8; Mismatches 32; Indels 28; Gaps 6;

QY 12 MAQGAMLAQERRVPRAA-----EVPGAQGGQGGPRGEE-----EAP---RG 49
Db 1 MTEGAR-AADEVPRVLGAPPFGPAALVGASPSGAPGREAERGSELGVSPSPAARG 59

QY 50 VRMAVPLLRMEGAPAGPGGRTAACFCTSRCLSRPW 87
Db 60 AELGADEQRVP-IPA----LAAIVFFCLGQTRPRSW 92

RESULT 7
Q9UGS8 PRELIMINARY; PRT; 100 AA.
AC Q9UGS8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 13, Last sequence update)
DE DJ745C22.1 (Novel protein) (Fragment).
GN BK747E2.2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

---

```
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ramsay H.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL031596; CAB63070.1; -.
FT NON_TER 100
SQ SEQUENCE 100 AA; 10683 MW; 023ADF640B07801C CRC64;

Query Match 11.9%; Score 69; DB 4; Length 100;
Best Local Similarity 38.9%; Pred. No. 20;
Matches 21; Conservative 7; Mismatches 18; Indels 8; Gaps 2;

QY 26 PRAAEVPGAQGGQGGPRGEEAPRGVMA-----VPLLRMEGAPAGPGGRTA 72
Db 3 PRSGGRPGATGRRRRRLRR-PRGLRCSRLPPPPPLPLLGLLLAAAGPGAARA 55

RESULT 8
Q8SOY5 PRELIMINARY; PRT; 102 AA.
AC Q8SOY5;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE P0485B12.32 protein (P0031D02.5 protein).
DE P0485B12.32 OR P0031D02.5.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1, PAC
clone:P0485B12.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone:P0031D02.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003348; BAB86501.1; -.
DR Gramene; Q8SOY5; -.
SQ SEQUENCE 102 AA; 10904 MW; F2B10683B72C651C CRC64;

Query Match 11.8%; Score 68.5; DB 10; Length 102;
Best Local Similarity 28.6%; Pred. No. 23;
Matches 20; Conservative 5; Mismatches 24; Indels 21; Gaps 2;

QY 22 ERRVPAARVPGAQGGQGGPRGEEAPRGVMAVPLLRMEGAPAGPGGRTAACFCTSRC 81
Db 2 EVRRPAGDEAATQ-----KGRPAAAGARRATP-----ASCAASTAAC 40

QY 82 LSRRPWKRSW 91
Db 41 QRRRPARSCW 50

RESULT 9
Q8N493 PRELIMINARY; PRT; 86 AA.
AC Q8N493;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein (Fragment).
```





RT "The genome sequence of *Drosophila melanogaster*."

RL Science 287:2185-2195(2000).  
DR EMBL; AE003568; AAF50815.1; -.  
DR FlyBase; FBgn0031176; CG1678.  
SQ SEQUENCE 85 AA; 8360 MW; 0

SQ SEQUENCE 85 AA; 8360 MW; 0F04923B4AEB66CE CRC64;

Query Match 11.1%; Score 64.5; DB 5; Length 85;

Best Local Similarity 29.5%; Pred. No. 47;

Matches	23	Conservative	12	Mismatches	28	Indels	15	Gaps	4
<p> </p>									

QY 8 LAF LMAQ GAMLAAQERRVPRAAEVPGAQQQGP--RGREEAPRGVRMAVPLLRMEGAP 64

[illegible]

Db 6 LIVLAMVALAAVQGRPPGL-----PQGLGLPQSGGLQPPQGI PQGP-----QGLP 55

QY 65 AGP--GGRTAACFSCTSR 80

**F**

Search completed: October 7, 2003, 14:02:37  
Job time : 97 secs

**THIS PAGE BLANK (USPTO)**

```
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- MISCELLANEOUS: X'S HAVE BEEN PLACED AT POSITIONS 10-12 BY HOMOLOGY
CC WITH THE COMPLETE SEQUENCE OF THE BOVINE PRECURSOR.
CC -!- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
DR HSSP: P01180; 2B2.
DR InterPro: IPR000981; Neurhyp_horm.
DR Pfam: PF00220; hormones; 1.
DR Pfam: PF00184; hormones; 1.
DR PRINTS: PR00831; NEUROPHYSIN.
DR ProDom: PD001676; Neurhyp_horm; 1.
DR SMART: SM00003; NH; 1.
DR PROSITE: PS00264; NEUROHYPOPHYS_HORM; 1.
KW Hormone; Hypothalamus; Cleavage on pair of basic residues; Amidation;
KW Vasoconstrictor.
FT NON_TER 1 1 ARG-VASOPRESSIN.
FT PEPTIDE 1 9
FT PEPTIDE 13 107 NEUROPHYSIN 2.
FT DISULFID 1 6
FT DISULFID 22 66 BY SIMILARITY.
FT DISULFID 25 39 BY SIMILARITY.
FT DISULFID 33 56 BY SIMILARITY.
FT DISULFID 40 46 BY SIMILARITY.
FT DISULFID 73 85 BY SIMILARITY.
FT DISULFID 79 97 BY SIMILARITY.
FT DISULFID 86 91 BY SIMILARITY.
FT MOD_RES 9 9 AMIDATION.
FT NON_TER 107 107
SQ SEQUENCE 107 AA; 11287 MW; 1E48830286092277 CRC64;

Query Match 9.6%; Score 55.5; DB 1; Length 107;
Best Local Similarity 34.2%; Pred. No. 1.8e+02;
Matches 13; Conservative 5; Mismatches 19; Indels 1; Gaps 1;

OY 44 EAPRGVGRMAVPLLRMEGAPAGPGGRTAACFSCTSRC 81
: ||| : : : ||| : : :
DB 4 QNCPGXXXXMSDLELRCLPCGPGGK-GRCFGPSICC 40

RESULT 15
REV_HV2NZ
ID REV_HV2NZ STANDARD; PRT; 107 AA.
AC P05873;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE REV protein (Anti-repression transactivator protein) (ART/TRS).
GN REV.
OS Human immunodeficiency virus type 2 (isolate NIH-2) (HIV-2).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11719;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88320359; PubMed=3261862;
RA Zagury J.F., Franchini G., Reitz M.S. Jr., Collalti E., Starcich B.R.,
RA Hall L., Fargnoli K., Jagodzinski L.L., Guo H.-G., Laure F.,
RA Arya S.K., Josephs S.F., Zagury D., Wong-Staal F., Gallo R.C.;
RT "Genetic variability between isolates of human immunodeficiency virus
RT (HIV) type 2 is comparable to the variability among HIV type 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:5941-5945(1988).
CC -!- FUNCTION: REV APPEARS TO ACT POST-TRANSCRIPTIONALLY TO RELIEVE
CC NEGATIVE REPRESSION OF GAG AND ENV PRODUCTION.
CC -!- SUBCELLULAR LOCATION: Nuclear; accumulates in the nucleoli.
CC -!- PTM: PHOSPHOPROTEIN WHOSE STATE OF PHOSPHORYLATION IS MEDIATED
CC BY A SPECIFIC SERINE KINASE ACTIVITY PRESENT IN THE NUCLEUS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC
```

```
DR EMBL: J03654; AAB00760.1; -.
DR HIV; J03654; REV$2NIH2.
DR InterPro: IPR000625; REV_protein.
DR Pfam: PF00424; REV; 1.
KW Transcription regulation; AIDS; Phosphorylation; Nuclear protein.
SQ SEQUENCE 107 AA; 12374 MW; 7B2341DF1AA304C1 CRC64;

Query Match 9.6%; Score 55.5; DB 1; Length 107;
Best Local Similarity 27.8%; Pred. No. 1.8e+02;
Matches 20; Conservative 7; Mismatches 30; Indels 15; Gaps 2;

OY 41 RGREAPRGVGRMAVPLLRMEGAPAGPGGRTAACFSCTSRCLSRRPWKRSW-----S 92
: ||| : : : ||| : : :
DB 4 RADEGLQRLRLRLHLLHQTNPYPQPG-----TASQRRRRRRKQWRQILALADS 56

OY 93 AGSCPMPHLS 104
: | | | |
DB 57 IYTFPPPADSP 68

Search completed: October 7, 2003, 14:00:55
Job time : 24 secs
```



```

RESULT 12
YC02_PARDE STANDARD; PRT; 89 AA.
AC P08302;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 01-AUG-1988 (Rel. 08, Last annotation update)
DE COX locus hypothetical protein 2.
OS Paracoccus denitrificans.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC Rhodobacteraceae; Paracoccus.
OX NCBI_TaxID=266;
RN [1]
SEQUENCE FROM N.A.
RA Raitio M., Jalli T., Saraste M.;
RT "Isolation and analysis of the genes for cytochrome c oxidase in
RT Paracoccus denitrificans.";
RL EMBO J. 6:2825-2833(1987).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X05828; CAA29270.1; -
DR PIR; S03805;
KW Hypothetical protein.
SQ SEQUENCE 89 AA; 9559 MW; FCC07D2BA9F026A CRC64;
Query Match 9.6%; Score 55.5; DB 1; Length 89;
Best Local Similarity 30.3%; Pred. No. 1.5e+02;
Matches 23; Conservative 8; Mismatches 26; Indels 19; Gaps 4;
QY 1 MLMAQALALF-----LMAQAMLAQAQRRVPRAAEVFGAGCGQGRREAPRGVRNAV 55
Db 23 VLLAFVALVGLSVKITQGMQGYDHR-PRASMLP-----PDPPAPNAAPVA 74
QY 56 LLRRMEGAPAGGGRT 71
Db 75 -----GTPAATGV 84

RESULT 13
DPD4_MOUSE STANDARD; PRT; 107 AA.
AC Q9CWF8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA polymerase delta subunit 4 (DNA polymerase delta subunit p12).
GN POLD4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito K.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,

```

```

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1- FUNCTION: The function of subunit 4 is not yet known.
CC -1- SUBUNIT: HETEROTETRAMER COMPOSED OF SUBUNITS OF 125 kDa, 50 kDa,
CC 66 kDa and 12 kDa (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE DELTA SUBUNIT 4 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AK010477; BAB26970.1; -
DR MGD; MGI:1916995; 2410012M21Rik.
DR Pfam; PF04081; DNA_pol_delta_4; 1.
KW DNA-directed DNA polymerase; DNA replication; Nuclear protein.
SQ SEQUENCE 107 AA; 12403 MW; 14A27054D4E7511E CRC64;
Query Match 9.6%; Score 55.5; DB 1; Length 107;
Best Local Similarity 24.6%; Pred. No. 1.8e+02;
Matches 17; Conservative 13; Mismatches 28; Indels 11; Gaps 2;
QY 42 GREAPRGVRMAYPLLRMEGAPAGPGGRTAACFSCRSCLSR-----RPMKRSWSA 93
Db 2 GRK---RFITDTPVVKRREGPPGHSGELAPELGDTLSQSETELELLRQDLAWQY 58
QY 94 GSCPGMPHL 102
Db 59 GPCTGTRL 67

RESULT 14
NEU2_BALPH STANDARD; PRT; 107 AA.
AC P01184;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Vasopressin-neurophysin 2 precursor (Fragment).
GN AVP.
OS Balaenoptera physalus (Finback whale) (Common rorqual).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Mysticeti;
OC Balaenopteridae; Balaenoptera.
OX NCBI_TaxID=9770;
RN [1]
SEQUENCE OF 1-9.
RA Acher R., Chauvet J., Chauvet M.-T.;
RT "Isolation of finback whale oxytocin and vasopressin.";
RL Nature 201:191-192(1964).
RN [2]
SEQUENCE OF 13-107.
RX MEDLINE=78148313; PubMed=639997;
RA Chauvet M.-T., Codogno P., Chauvet J., Acher R.;
RT "Phylogeny of neurophysins: complete amino acid sequence of whale
RT (Balaenoptera physalus) MSEL-neurophysin.";
RL FEBS Lett. 88:91-93(1978).
CC -1- FUNCTION: Neurophysin 2 specifically binds vasopressin.
CC -1- FUNCTION: Vasopressin has a direct antidiuretic action on the
CC kidney, it also causes vasoconstriction of the peripheral vessels.

```

```
[1]
RN  SEQUENCE FROM N.A.
RP  STRAIN=cv. Neepawa, and cv. Yamhill; TISSUE=Embryo;
RC  MEDLINE=93049355; PubMed=1425704;
RA  Kawashima I., Kennedy T.D., Chino M., Lane B.G.;
RT  "Wheat Ec metallothionein genes. Like mammalian Zn2+ metallothionein
RT  genes, wheat Zn2+ metallothionein genes are conspicuously expressed
RT  during embryogenesis.";
RL  Eur. J. Biochem. 209:971-976(1992).
CC  -|- FUNCTION: BINDS 5 MOLECULES OF ZINC. MAY HAVE A ROLE IN ZN(2+)
CC  HOMEOSTASIS DURING EMBRYOGENESIS.
CC  -|- DEVELOPMENTAL STAGE: EXPRESSED DURING EMBRYOGENESIS.
CC  -|- INDUCTION: BY abscisic acid (ABA).
CC  -|- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 15.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; X68291; CAA48351.1; -
DR  EMBL; X68288; CAA48348.1; -
DR  PIR; S27366; S27366.
DR  InterPro; IPR000316; Metallothion_15.
DR  Pfam; PF02068; Metallothio_PEC; 1.
DR  PRINTS; PR00877; MTPLANTPEC.
DR  Metal-binding; Metal-thiolate cluster; Zinc.
DR  INIT_MET 0
FT  SEQUENCE 80 AA; 7668 MW; 832AFBCEDE8F2FAE CRC64;
SQ  SEQUENCE 9.7%; Score 56; DB 1; Length 80;
Query Match 28.6%; Pred. No. 1.3e+02;
Best Local Similarity 28.6%; Mismatches 4; Indels 10; Gaps 2;
Matches 18; Conservative 4;
QY 28 AAEVPCGCGTCTARSADNAGEHTTCGCGHCGNCPACGREGTFSGRANRANC-SCG 57
Db 9 AVPCGCGTCTARSADNAGEHTTCGCGHCGNCPACGREGTFSGRANRANC-SCG 57
QY 79 SRC 81
Db 68 AAC 70
RESULT 10
ID REV_HV2RO STANDARD; PRT; 100 AA.
AC P04615;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE REV protein (Anti-repression transactivator protein) (ART/TRS).
GN REV.
OS Human immunodeficiency virus type 2 (isolate ROD) (HIV-2).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11720;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87173056; PubMed=3031510;
RA Guyader M., Emerman M., Sonigo P., Clavel F., Montagnier L.,
RA Alizon M.;
RT "Genome organization and transactivation of the human
RT Immunodeficiency virus type 2.";
RL Nature 326:662-669(1987).
CC -|- FUNCTION: REV APPEARS TO ACT POST-TRANSCRIPTIONALLY TO RELIEVE
CC NEGATIVE REPRESSION OF GAG AND ENV PRODUCTION.
CC -|- SUBCELLULAR LOCATION: Nuclear; accumulates in the nucleoli.
CC -|- PTM: PHOSPHOPROTEIN WHOSE STATE OF PHOSPHORYLATION IS MEDIATED
CC BY A SPECIFIC SERINE KINASE ACTIVITY PRESENT IN THE NUCLEUS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR  EMBL; X68291; CAA48351.1; -
DR  EMBL; X68288; CAA48348.1; -
DR  PIR; S27366; S27366.
DR  InterPro; IPR000316; Metallothion_15.
DR  Pfam; PF02068; Metallothio_PEC; 1.
DR  PRINTS; PR00877; MTPLANTPEC.
DR  Metal-binding; Metal-thiolate cluster; Zinc.
DR  INIT_MET 0
FT  SEQUENCE 80 AA; 7668 MW; 832AFBCEDE8F2FAE CRC64;
SQ  SEQUENCE 9.7%; Score 56; DB 1; Length 80;
Query Match 28.6%; Pred. No. 1.3e+02;
Best Local Similarity 28.6%; Mismatches 4; Indels 10; Gaps 2;
Matches 18; Conservative 4;
QY 28 AAEVPCGCGTCTARSADNAGEHTTCGCGHCGNCPACGREGTFSGRANRANC-SCG 57
Db 9 AVPCGCGTCTARSADNAGEHTTCGCGHCGNCPACGREGTFSGRANRANC-SCG 57
QY 79 SRC 81
Db 68 AAC 70
RESULT 10
ID REV_HV2RO STANDARD; PRT; 100 AA.
AC P04615;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE REV protein (Anti-repression transactivator protein) (ART/TRS).
GN REV.
OS Human immunodeficiency virus type 2 (isolate ROD) (HIV-2).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11720;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87173056; PubMed=3031510;
RA Guyader M., Emerman M., Sonigo P., Clavel F., Montagnier L.,
RA Alizon M.;
RT "Genome organization and transactivation of the human
RT Immunodeficiency virus type 2.";
RL Nature 326:662-669(1987).
CC -|- FUNCTION: REV APPEARS TO ACT POST-TRANSCRIPTIONALLY TO RELIEVE
CC NEGATIVE REPRESSION OF GAG AND ENV PRODUCTION.
CC -|- SUBCELLULAR LOCATION: Nuclear; accumulates in the nucleoli.
CC -|- PTM: PHOSPHOPROTEIN WHOSE STATE OF PHOSPHORYLATION IS MEDIATED
CC BY A SPECIFIC SERINE KINASE ACTIVITY PRESENT IN THE NUCLEUS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR  EMBL; X68291; CAA48351.1; -
DR  EMBL; X68288; CAA48348.1; -
DR  PIR; S27366; S27366.
DR  InterPro; IPR000316; Metallothion_15.
DR  Pfam; PF02068; Metallothio_PEC; 1.
DR  PRINTS; PR00877; MTPLANTPEC.
DR  Metal-binding; Metal-thiolate cluster; Zinc.
DR  INIT_MET 0
FT  SEQUENCE 80 AA; 7668 MW; 832AFBCEDE8F2FAE CRC64;
SQ  SEQUENCE 9.7%; Score 56; DB 1; Length 100;
Query Match 29.3%; Pred. No. 1.5e+02;
Best Local Similarity 29.3%; Mismatches 7; Indels 16; Gaps 3;
Matches 22; Conservative 7;
QY 41 RGREAPRGVMAVPLLRMEGAPAGPGGRTAACFCSTSCLSRRPWKRSW-
Db 4 RADEEGLRKURLRLHLQNTNPYQGG-
QY 93 AGSCPGMPLSP-DQ 106
Db 57 IYTFDPADSPLDQ 71
RESULT 11
ID YDH3_HSVSC STANDARD; PRT; 103 AA.
AC P22577;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 9.5 kDa protein in DHFR 3'region (ORF3).
OS Herpesvirus saimiri (subgroup C / strain 488).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
OX NCBI_TaxID=10384;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90266466; PubMed=2161148;
RA Biesinger B., Trimble J.J., Desrosiers R.C., Fleckenstein B.;
RT "The divergence between two oncogenic Herpesvirus saimiri strains in
RT a genomic region related to the transforming phenotype.";
RL Virology 176:505-514(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR  EMBL; M5264; AAA72930.1; -
DR  InterPro; IPR000087; Collagen.
KW Hypothetical protein.
SQ SEQUENCE 103 AA; 9457 MW; 1FC2E48FF239D917 CRC64;
Query Match 9.7%; Score 56; DB 1; Length 103;
Best Local Similarity 28.6%; Pred. No. 1.6e+02;
Matches 20; Conservative 1; Mismatches 27; Indels 22; Gaps 3;
QY 32 PCAQQQQQPRGEEAPRGVMAVPLLRMEGAPAGPGGRTAACFCSTSCLSRRPWKRSW 91
Db 12 PGSGPGGPGGPGGPGGPGG-
QY 92 SAGSC-PGMP 100
Db 51 PGSGPGGPGGPG 60
```



RA Chakrabarti L., Guyader M., Alison M., Daniel M.D., Desrosiers R.C.,  
 RA Tiollais P., Sonigo P.;  
 RT "Sequence of simian immunodeficiency virus from macaque and its  
 RT relationship to other human and simian retroviruses.";  
 RL Nature 328:543-547(1987).  
 CC -!- FUNCTION: REV APPEARS TO ACT POST-TRANSCRIPTIONALLY TO RELIEVE  
 CC NEGATIVE REPRESSION OF GAG AND ENV PRODUCTION.  
 CC -!- SUBCELLULAR LOCATION: Nuclear; accumulates in the nucleoli.  
 CC -!- PTM: PHOSPHOPROTEIN WHOSE STATE OF PHOSPHORYLATION IS MEDIATED  
 CC BY A SPECIFIC SERINE KINASE ACTIVITY PRESENT IN THE NUCLEUS.  
 CC -!- MISCELLANEOUS: THIS IS A MACAQUE ISOLATE.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: Y00277; CAB46522.1; ..  
 DR HIV: M16403; REV\$MM142.  
 DR InterPro: IPR000625; REV\_protein.  
 DR Pfam: PF00424; REV; 1.  
 KW Transcription regulation; AIDS; Phosphorylation; Nuclear protein.  
 SQ SEQUENCE 107 AA; 12443 MW; 8E30BFD94ABA50B8 CRC64;  
 Query Match 10.2%; Score 59; DB 1; Length 107;  
 Best Local Similarity 32.0%; Pred. No. 91;  
 Matches 16; Conservative 6; Mismatches 20; Indels 8; Gaps 2;  
 QY 42 GREEAPRGVMAVPLLRMEGAPAGPGGRTAACFCTSCLSRRPWRKSW 91  
 DB 6 GEEELRRRLRL-IHLHQTNPTGSG-----SANQRQRKRRWRQW 47  
 RESULT 5  
 CALL\_RABIT STANDARD; PRT; 53 AA.  
 AC P02456;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE Collagen alpha 1(I) chain (Fragment).  
 GN COL1A1.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=70252720; PubMed=4194291;  
 RA Bornstein P., Nesse R.;  
 RT "The comparative biochemistry of collagen: the structure of rabbit  
 RT skin collagen and its relevance to immunochemical studies of  
 RT collagen.";  
 RL Arch. Biochem. Biophys. 138:443-450(1970).  
 CC -!- FUNCTION: TYPE I COLLAGEN IS A MEMBER OF GROUP I COLLAGEN  
 CC (FIBRILLAR FORMING COLLAGEN).  
 CC -!- SUBUNIT: TRIMERS OF ONE ALPHA 2(I) AND TWO ALPHA 1(I) CHAINS.  
 CC -!- TISSUE SPECIFICITY: FORMS THE FIBRILS OF TENDON, LIGAMENTS AND  
 CC BONES. IN BONES THE FIBRILS ARE MINERALIZED WITH CALCIUM  
 CC HYDROXYAPATITE.  
 CC -!- PTM: Prolines at the third position of the tripeptide repeating  
 CC unit (G-X-Y) are hydroxylated in some or all of the chains.  
 DR PIR: A02856; CGBLIS.  
 DR InterPro: IPR000087; Collagen.  
 DR InterPro: IPR001007; VWF\_C.  
 DR Pfam: PF01391; Collagen; 1.  
 DR PROSITE: PS01208; VWFC\_1; PARTIAL.  
 KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;  
 KW Collagen.  
 FT MOD\_RES 7 7 CONVERTED TO AN ALDEHYDE GROUP THAT IS

FT MOD\_RES 26 INVOLVED IN CROSS-LINKING.  
 FT MOD\_RES 29 HYDROXYLATION (PROBABLE).  
 FT MOD\_RES 32 HYDROXYLATION (PROBABLE).  
 FT MOD\_RES 41 HYDROXYLATION (PROBABLE).  
 FT MOD\_RES 44 HYDROXYLATION (PROBABLE).  
 FT MOD\_RES 47 HYDROXYLATION (PROBABLE).  
 FT MOD\_RES 53  
 SQ SEQUENCE 53 AA; 4987 MW; 127582E5E52B87FC CRC64;  
 Query Match 10.1%; Score 58.5; DB 1; Length 53;  
 Best Local Similarity 36.8%; Pred. No. 56;  
 Matches 14; Conservative 1; Mismatches 18; Indels 5; Gaps 1;  
 QY 31 VPGAGQGQPGRGREAPRGVMAVPLLRMEGAPAGPG 68  
 DB 13 VPGMGPSGPRGLPGPP-----GAPGPZGFZGPPGZPG 45  
 RESULT 6  
 CALC\_BOVIN STANDARD; PRT; 86 AA.  
 ID CALC\_BOVIN  
 AC P25508;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Collagen alpha 1(XII) chain (Fragments).  
 GN COL12A1.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=88255287; PubMed=3133242;  
 RA Dublet B., Dixon E., de Miguel E., van der Rest M.;  
 RT "Bovine type XII collagen: amino acid sequence of a 10 kDa pepsin  
 RT fragment from periodontal ligament reveals a high degree of homology  
 RT with the chicken alpha 1(XII) sequence.";  
 RL FEBS Lett. 233:177-180(1988).  
 CC -!- FUNCTION: TYPE XII COLLAGEN INTERACTS WITH TYPE I COLLAGEN-  
 CC CONTAINING FIBRILS, THE COL1 DOMAIN COULD BE ASSOCIATED WITH THE  
 CC SURFACE OF THE FIBRILS, AND THE COL2 AND NC3 DOMAINS MAY BE  
 CC LOCALIZED IN THE PERIFIBRILLAR MATRIX.  
 CC -!- SUBUNIT: TRIMER OF IDENTICAL CHAINS EACH CONTAINING 190 kDa OF  
 CC NONTRIPLE-HELICAL SEQUENCES.  
 CC -!- PTM: The triple-helical tail is stabilized by disulfide bonds at  
 CC each end.  
 CC -!- PTM: Prolines at the third position of the tripeptide repeating  
 CC unit (G-X-Y) are hydroxylated in some or all of the chains.  
 CC -!- SIMILARITY: HIGH TO COLLAGEN ALPHA 1(IX).  
 DR PIR: S00802; S00802.  
 DR InterPro: IPR000087; Collagen.  
 DR Extracellular matrix; Connective tissue; Repeat; Hydroxylation;  
 KW Cell adhesion; Collagen.  
 FT NON\_TER 1 1  
 FT NON\_CONS 15 16  
 FT NON\_CONS 48 49  
 FT MOD\_RES 6 6 HYDROXYLATION.  
 FT MOD\_RES 9 9 HYDROXYLATION.  
 FT MOD\_RES 12 12 HYDROXYLATION.  
 FT MOD\_RES 18 18 HYDROXYLATION.  
 FT MOD\_RES 24 24 HYDROXYLATION.  
 FT MOD\_RES 27 27 HYDROXYLATION.  
 FT MOD\_RES 30 30 HYDROXYLATION.  
 FT MOD\_RES 42 42 HYDROXYLATION.  
 FT MOD\_RES 51 51 HYDROXYLATION.  
 FT MOD\_RES 54 54 HYDROXYLATION.  
 FT MOD\_RES 65 65 HYDROXYLATION.  
 FT MOD\_RES 74 74 HYDROXYLATION.  
 FT MOD\_RES 77 77 HYDROXYLATION.  
 FT MOD\_RES 80 80 HYDROXYLATION.

```
REV_HV2ST
ID REV_HV2ST STANDARD; PRT; 107 AA.
AC P20870;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE REV protein (Anti-repression transactivator protein) (ART/TRS).
GN REV.
OS Human immunodeficiency virus type 2 (isolate ST) (HIV-2).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11721;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90112662; PubMed=2296086;
RA Kumar P., Hui H., Kappes J.C., Haggarty B.S., Hoxie J.A., Arya S.K.,
RA Shaw G.M., Hahn B.H.;
RT "Molecular characterization of an attenuated human immunodeficiency
RT virus type 2 isolate."
RL J. Virol. 64:890-901(1990).
CC -1- FUNCTION: REV APPEARS TO ACT POST-TRANSCRIPTIONALLY TO RELIEVE
CC -1- NEGATIVE REPRESSION OF GAG AND ENV PRODUCTION.
CC -1- SUBCELLULAR LOCATION: Nuclear; accumulates in the nucleoli.
CC -1- PTM: PHOSPHOPROTEIN WHOSE STATE OF PHOSPHORYLATION IS MEDIATED
CC BY A SPECIFIC SERINE KINASE ACTIVITY PRESENT IN THE NUCLEUS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M31113; AAB01357.1; -
DR PIR; G33943; VKLJST.
DR HIV; M31113; REV32ST.
DR InterPro; IPR000625; REV_protein.
DR Pfam; PF00424; REV; 1.
KW Transcription regulation; AIDS; Phosphorylation; Nuclear protein.
SQ SEQUENCE 107 AA; 12576 MW; AFF998FD8C55EA58 CRC64;
Query Match 10.3%; Score 59.5; DB 1; Length 107;
Best Local Similarity 31.4%; Pred. No. 83;
Matches 16; Conservative 5; Mismatches 23; Indels 7; Gaps 1;
QY 41 RGEEAPRGVRMAVPLLRMEGAPAGPGGRTAACFCTSRCLSRPWRKSW 91
Db 4 RAEEELRRKRLRLILLHOTNPYPQPG-----TASQRNRRRRWKQW 47
RESULT 3
PRK1_HUMAN STANDARD; PRT; 105 AA.
AC P58294;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Prokineticin 1 precursor (Endocrine-gland-derived vascular endothelial
DE growth factor) (EG-VEGF) (Mambakine).
GN PROK1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21160229; PubMed=11259612;
RA Li M., Bullock C.M., Knauer D.J., Ehler F.J., Zhou Q.Y.;
RT "Identification of two prokineticin cDNAs: recombinant proteins
RT potentially contract gastrointestinal smooth muscle."
RL Mol. Pharmacol. 59:692-698(2001).
RN [2]
RP SEQUENCE FROM N.A.
```

```
RX MEDLINE=21419730; PubMed=11528470;
RA LeCouter J., Kowalski J., Foster J., Hass P., Zhang Z.,
RA Dillard-Telm L., Frantz G., Rangell L., Decuzman L., Keller G.-A.,
RA Peale F., Gurney A., Hillan K.J., Ferrara N.;
RT "Identification of an angiogenic mitogen selective for endocrine gland
RT endothelium."
RL Nature 412:877-884(2001).
RN [3]
RP SEQUENCE FROM N.A.
RA Fraser C.;
RT "Mambakine, a snake venom related endocrine hormone that controls
RT macrophages."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Potently contract gastrointestinal (GI) smooth muscle.
CC Induces proliferation, migration and fenestration (the formation
CC of membrane discontinuities) in capillary endothelial cells
CC derived from endocrine glands. Has little or no effect on a
CC variety of other endothelial and non-endothelial cell types.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed in the steroidogenic glands, ovary,
CC testis, adrenal and placenta.
CC -1- SIMILARITY: BELONGS TO THE PROKINETICIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF333024; AAK49918.1; -
DR EMBL; AY029225; AAK33111.1; -
DR Genew; HGNC:18454; PROK1.
DR MIM; 606233; -
DR Mitogen; Growth factor; Signal.
FT SIGNAL 1 19 BY SIMILARITY.
FT CHAIN 20 105 PROKINETICIN 1.
FT DISULFID 26 38 BY SIMILARITY.
FT DISULFID 32 50 BY SIMILARITY.
FT DISULFID 37 78 BY SIMILARITY.
FT DISULFID 60 86 BY SIMILARITY.
FT DISULFID 80 96 BY SIMILARITY.
SQ SEQUENCE 105 AA; 11715 MW; C7E3FDE30EFB416A CRC64;
Query Match 10.2%; Score 59; DB 1; Length 105;
Best Local Similarity 31.8%; Pred. No. 90;
Matches 21; Conservative 7; Mismatches 18; Indels 20; Gaps 4;
QY 48 RGVMAVPLLRMEGAPAGPGGRTAACFCTSRCLSRPWRKSWAGSCPGMPLH----S 103
Db 45 RGLRMCTPLGR--EGECHPGSH-----KVPFPRKRKHTCPLNLLCSRF 89
QY 104 PDGGRF 109
Db 90 PD-GRY 94
RESULT 4
REV_SIVM1 STANDARD; PRT; 107 AA.
AC P05875; Q85724;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE REV protein (Anti-repression transactivator protein) (ART/TRS).
GN REV.
OS Simian immunodeficiency virus (Mml42-83 isolate) (SIV-MAC).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11733;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87287230; PubMed=3649576;
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:34:13 ; Search time 22 Seconds  
(without alignments)  
232.996 Million cell updates/sec

Title: US-09-807-512-2  
Perfect score: 579  
Sequence: 1 MLMQAEALAFMAQAMLA.....SWSAGSCPGMPLSPDQGRF 109

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 16345

Minimum DB seq length: 0  
Maximum DB seq length: 109

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66.5	11.5	74	1 M84B_DROME	Q01643 drosophila
2	59.5	10.3	107	1 REV_HV2ST	P20870 human immun
3	59	10.2	105	1 PRK1_HUMAN	P58294 homo sapien
4	59	10.2	107	1 REV_SIVM1	P05875 simian immu
5	58.5	10.1	53	1 CAL1_RABIT	P02456 oryctolagus
6	58.5	10.1	86	1 CA1C_BOVIN	P25508 bos taurus
7	58.5	10.1	108	1 REV_SIVM2	P05874 simian immu
8	56.5	9.8	82	1 PN3I_PENVA	Q963c5 penaeus van
9	56	9.7	80	1 EC3_WHEAT	P30570 triticum ae
10	56	9.7	100	1 REV_HV2RO	P04615 human immun
11	56	9.7	103	1 YDH3_HSVSC	P22577 herpesvirus
12	55.5	9.6	89	1 YCO2_PARDE	P08302 paracoccus
13	55.5	9.6	107	1 DPD4_MOUSE	Q9cwp8 mus musculu
14	55.5	9.6	107	1 NEU2_BALPH	P01184 balaenopter
15	55.5	9.6	107	1 REV_HV2NZ	P05873 human immun
16	55	9.5	100	1 REV_SIVSP	P19502 simian immu
17	54.5	9.4	78	1 DRPH_CARMA	Q06202 carcarius ma
18	54	9.3	79	1 ET2_MACFA	Q28470 macaca fasc
19	54	9.3	105	1 IBA2_HUMAN	Q14657 homo sapien
20	52.5	9.1	100	1 CHA3_BOMMO	P08929 bombyx mori
21	52.5	9.1	109	1 YW73_MYCTU	Q50690 mycobacteri
22	52	9.0	80	1 ECL1_WHEAT	P30569 triticum ae
23	52	9.0	92	1 SY04_HUMAN	P13236 h small ind
24	52	9.0	92	1 SY04_MOUSE	P14097 mus musculu
25	52	9.0	109	1 HMGC_HUMAN	P52926 homo sapien
26	51.5	8.9	83	1 VIE3_AGR75	P08061 agrobacteri
27	51.5	8.9	100	1 REV_HV2CA	P24104 human immun
28	51.5	8.9	101	1 REV_SIVS4	P12486 simian immu
29	50.5	8.7	62	1 SCX4_TITFA	P83435 titiyus fasc
30	50.5	8.7	62	1 TTX6_TITSE	P45669 titiyus serr
31	50.5	8.7	84	1 NTXP_TITSE	O77463 titiyus serr
32	50.5	8.7	101	1 RS14_BRUME	O8yhm7 bruceella me
33	50.5	8.7	102	1 COLL_HSVSC	P22576 herpesvirus

RESULT 2

34 50 8.6 92 1 SY04\_RAT P50230 rattus norv  
35 50 8.6 100 1 DEF6\_HUMAN Q01524 homo sapien  
36 50 8.6 100 1 KRF2\_COLLI O93499 columba liv  
37 50 8.6 100 1 KRF3\_COLLI O93500 columba liv  
38 50 8.6 109 1 2M33\_MAIZE O82106 zea mays (m  
39 49.5 8.5 82 1 PN3B\_PENVA P81039 penaeus van  
40 49.5 8.5 85 1 NEUL\_PAPHA P32005 papio hamad  
41 49.5 8.5 93 1 DEF1\_CAVPO P11478 cavia porce  
42 49.5 8.5 105 1 COLL\_HSVS7 P25050 herpesvirus  
43 49 8.5 73 1 BP4C\_BRANA P41506 brassica na  
44 49 8.5 92 1 SY04\_RABIT P46632 oryctolagus  
45 49 8.5 108 1 HMGC\_MOUSE P25927 mus musculu

## ALIGNMENTS

RESULT 1  
M84B\_DROME  
ID M84B\_DROME STANDARD; PRT; 74 AA.  
AC O01643;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Male specific sperm protein Mst84Db.  
GN Mst84Db.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Oregon-R;  
RX MEDLINE=92102953; PubMed=1684716;  
RA Kuhn R., Kuhn C., Boersch D., Glaetzer K.H., Schaefer U.,  
RA Schaefer M.;  
RT "A cluster of four genes selectively expressed in the male germ line  
of Drosophila melanogaster.";  
RL Mech. Dev. 35:143-151(1991).  
CC -!- TISSUE SPECIFICITY: Testis.  
CC -!- DEVELOPMENTAL STAGE: PRIMARY SPERMATOCYTES.  
CC -!- DOMAIN: THIS PROTEIN IS MOSTLY COMPOSED OF REPETITIVE C-G-P  
MOTIFS.  
CC -!- SIMILARITY: BELONGS TO THE MST(3)CGP FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL; X67703; CAA47938.1; -.  
CC PIR; S25773; S25773.  
DR HSSP; P01180; INPO.  
DR FlyBase; FBgn0004173; Mst84Db.  
DR InterPro; IPR005634; MSSP.  
DR Pfam; PF03940; MSSP; 1  
DR Spermatogenesis; Repeat; Multigene family.  
KW Spermatogenesis; Repeat; Multigene family.  
SQ SEQUENCE 74 AA; 6824 MW; D0793137A7E7D1D0 CRC64;

Query Match 11.5%; Score 66.5; DB 1; Length 74;  
Best Local Similarity 42.1%; Pred. No. 16;  
Matches 16; Conservative 0; Mismatches 21; Indels 1; Gaps 1;

QY 62 GAPAGCGRTAAFCSTSRCLS--RRPWKRWSAGSCPG 98

Db 16 GPGCGCGCGCGSCCPCGCCAPWGPCGCGCCG 53

C;Keywords: coiled coil; extracellular matrix; glycoprotein; trimer; triple helix  
F;7/Modified site: allylsine (Lys) #status predicted

Query Match 10.1%; Score 58.5; DB 1; Length 53;

Best Local Similarity 36.8%; Pred. No. 79;

Matches 14; Conservative 1; Mismatches 18; Indels 5; Gaps 1;

QY 31 VPGAQGGQGRGEEAPRGVMAVPLLRMEGAPAGPG 68

Db 13 VPGMGPSGRGLGPGP-----GAPGPGZGFZGPGZPG 45

#### RESULT 13

Q00804

core protein - hepatitis C virus (fragment)

C;Species: hepatitis C virus

C;Date: 03-May-1994 #sequence\_revision 07-Oct-1994 #text\_change 17-Nov-2000

C;Accession: Q00804; PQ0803

R;Simmonds, P.; McOmish, F.; Yap, P.L.; Chan, S.W.; Lin, C.K.; Dusheiko, G.; Saeed, A.A.

J. gen. Virol. 74, 661-668, 1993

A;Title: Sequence variability in the 5' non-coding region of hepatitis C virus: identification

A;Reference number: PQ0803; MUID:93224886; PMID:8385694

A;Accession: PQ0804

A;Molecule type: mRNA

A;Residues: 1-82 <SIM>

A;Experimental source: variant EG-33 type 4

A;Accession: PQ0803

A;Molecule type: mRNA

A;Residues: 1-82 <SI2>

A;Experimental source: variant EG-29 type 4

C;Superfamily: hepatitis C virus genome polyprotein

C;Keywords: capsid protein; core protein

Query Match 10.1%; Score 58.5; DB 2; Length 82;

Best Local Similarity 25.0%; Pred. No. 1.2e+02;

Matches 24; Conservative 4; Mismatches 19; Indels 49; Gaps 5;

QY 24 RVPRAAEVFGAQGQ-----QGPR-----GREAPRGVMAVPLLRMEGAP 64

Db 10 RRPMDVKFPFG-QQIVGGVYLLPRGRLGVRAKTRKTSERSQPRGRQPIPKARRSEG-- 66

QY 65 AGPGGRTAACFSCSCLSRPWKRSWSAGSCPGMP 100

Db 67 -----RSWAQ---PGYP 75

#### RESULT 14

S00802

collagen alpha 1(XII) chain - bovine (fragments)

C;Species: Bos primigenius taurus (cattle)

C;Date: 01-Dec-1989 #sequence\_revision 01-Dec-1989 #text\_change 08-Dec-1995

C;Accession: S00802

R;Dublet, B.; Dixon, E.; de Miguel, E.; van der Rest, M.

FEBS Lett. 233, 177-180, 1988

A;Title: Bovine type XII collagen: amino acid sequence of a 10 kDa pepsin fragment from

A;Reference number: S00802; MUID:88255287; PMID:3133242

A;Accession: S00802

A;Molecule type: protein

A;Residues: 1-15;16-48;49-86 <DUB>

C;Superfamily: collagen alpha 1(XII) chain; fibronectin type III repeat homology; von Wi

C;Keywords: cell binding; coiled coil; extracellular matrix; glycoprotein; hydroxyprolin

F;6,9,12,18,24,27,30,42,51,54,65,74,77,80/Modified site: hydroxyproline (Pro) #status ex

Query Match

Best Local Similarity 28.4%; Score 58.5; DB 2; Length 86;

Matches 25; Conservative 5; Mismatches 31; Indels 27; Gaps 4;

QY 32 PGAGQGGQGRGEEAPRGVMAVPLLRMEGAPAGPG-----GRTAACFSCTS 79

Db 12 PGAGGPGGGRGPFPGTGMQPGQGER-GLPGXGGERGLPGPPGQGESRTGPGSGTGS 70

QY 80 RCLRRRPWKRWSAGSCPGMPLSPDQG 107

Db 80 RCLRRRPWKRWSAGSCPGMPLSPDQG 107

Db 71 R-----GPPGPPGRPGDSG 84

#### RESULT 15

JQ2283

negatively phytochrome regulated protein II - swollen duckweed

N;Alternate names: NRP2

C;Species: Lemna gibba (swollen duckweed)

C;Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 21-Jul-2000

C;Accession: JQ2283; SI8967

R;Okubara, P.A.; Williams, S.A.; Doxsee, R.A.; Tobin, E.M.

Plant Physiol. 101, 915-924, 1993

A;Title: Analysis of genes negatively regulated by phytochrome action in Lemna gibba

A;Reference number: JQ2282; MUID:94143480; PMID:8310060

A;Accession: JQ2283

A;Molecule type: DNA

A;Residues: 1-103 <OKU>

A;Cross-references: EMBL:X64145; MID:gl9405; PIDN:CAA45506.1; PID:gl9406

A;Experimental source: cv. G-3

C;Genetics:

A;Gene: NPR2

Query Match 10.1%; Score 58.5; DB 2; Length 103;

Best Local Similarity 35.6%; Pred. No. 1.4e+02;

Matches 21; Conservative 6; Mismatches 23; Indels 9; Gaps 2;

QY 19 AAQERRVPRAAEVPGAQGGQ-----PRGEEAPRGVMAVPLLRMEGAPAGPG 69

Db 43 AAQKEQA-RAATVPGTRRRGTAARWTPAPSARPPAPGPGHGAQPAERWDPDWSTPG 100

Search completed: October 7, 2003, 14:03:58

Job time : 41 secs

Db 54 PGDQGNPGYKPGIP-----GPTGPDGAQGAKGSC----- 84

QY 92 SAGSCPGMPLSP 104  
Db 85 --DHCP-TPRLFP 94

## RESULT 8

I45876

collagen alpha 1(II) chain - bovine (fragment)

C:Species: Bos primigenius taurus (cattle)

C:Date: 15-Oct-1996 #sequence\_revision 15-Oct-1996 #text\_change 13-Aug-1999

C:Accession: I45876

R:Brand, D.D.; Myers, L.K.; Terato, K.; Whittington, K.B.; Stuart, J.M.; Rosloniec, E.F.

J. Immunol. 152, 3088-3097, 1994

A:Title: Characterization of the T cell determinants in the induction of autoimmune arthritis

A:Reference number: I45876; MUID:94194070; PMID:7511638

A:Accession: I45876

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-93 &lt;BRA&gt;

A:Cross-references: GB:L28918; NID:9457186; PIDN:AAA30436.1; PID:9457187

C:Genetics:

A:Gene: CB11

C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;

Query Match 10.4%; Score 60.5; DB 2; Length 93;

Best Local Similarity 38.3%; Pred. No. 85;

Matches 18; Conservative 3; Mismatches 11; Indels 15; Gaps 3;

QY 32 PGAGQOQGP---RGRE--EAPRGVMAVPLLRMEGAPAGGRTAA 73

Db 7 PEAKGEAGTGARGPGAQGPGRG-----EPGTGSPGPAGAA 43

## RESULT 9

B61396

collagen alpha 1(IX) chain - golden hamster (fragment)

C:Species: Mesocricetus auratus (golden hamster)

C:Date: 09-Sep-1994 #sequence\_revision 09-Sep-1994 #text\_change 07-May-1999

C:Accession: B61396

R:Cizdziel, P.E.; Hosoi, J.; Montgomery, J.C.; Wiseman, R.W.; Barrett, J.C.

Mol. Carcinog. 4, 14-24, 1991

A:Title: Loss of a tumor suppressor gene function is correlated with downregulation of

A:Reference number: A61396; MUID:91182265; PMID:2009131

A:Accession: B61396

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-107 &lt;CI2&gt;

C:Superfamily: unassigned collagens

Query Match 10.4%; Score 60.5; DB 2; Length 107;

Best Local Similarity 34.2%; Pred. No. 96;

Matches 26; Conservative 4; Mismatches 27; Indels 19; Gaps 6;

QY 25 VPRAEVPGAQG--QQGPRGR--EAPRGVVR--MAVPLLRMEGAPAGGRTAAC--FS 76

Db 39 LPFAGLPGDPGPASYGKNGRDEQGPVGAGIPGV-----GPPGPPGPPGFCFEPAS 91

QY 77 CT----SRCLGRRRPWK 88

Db 92 CTLQSGQRAFSGKPEK 107

## RESULT 10

PQ0805

core protein - hepatitis C virus (fragment)

C:Species: hepatitis C virus

C:Date: 03-May-1994 #sequence\_revision 07-Oct-1994 #text\_change 02-Feb-2001

C:Accession: PQ0805

R:Simmonds, P.; McOmish, F.; Yap, P.L.; Chan, S.W.; Lin, C.K.; Dusheiko, G.; Saeed, A.A.

J. Gen. Virol. 74, 661-668, 1993

A:Title: Sequence variability in the 5' non-coding region of hepatitis C virus: ident

A:Reference number: PQ0803; MUID:93224886; PMID:8385694

A:Accession: PQ0805

A:Molecule type: mRNA

A:Residues: 1-75 <SIM>

A:Experimental source: variant, EG-21 type 4

C:Superfamily: hepatitis C virus genome polyprotein

C:Keywords: core protein; nucleotide binding; P-loop

F:34-41/Region: nucleotide-binding motif A (P-loop)

Query Match 10.3%; Score 59.5; DB 2; Length 75;

Best Local Similarity 25.0%; Pred. No. 87;

Matches 24; Conservative 4; Mismatches 19; Indels 49; Gaps 5;

QY 24 RVPRAEVPGAQGO-----OGPR-----GEEAPRGVMAVPLLRMEGAP 64

Db 6 RRPDVKRPGG-GQIVGVYLLPRGPRLGVRATKTSRSPGRGRQPIPAKRSQ-- 62

QY 65 AGPGGRTAACFSCTSRCLSRRPKRSWSAGSCPGMP 100

Db 63 -----RSWAQ---PGFP 71

## RESULT 11

VKLJST

trans-regulatory splicing protein - human immunodeficiency virus type 2 (isolate ST)

N:Alternate names: anti-repression trans-activator; art protein; rev protein; trs pro

C:Species: human immunodeficiency virus type 2, HIV-2

A:Note: host Homo sapiens (man)

C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 16-Jul-1999

C:Accession: G33943

R:Kumar, P.; Hui, H.; Kappes, J.C.; Hagarty, B.S.; Hoxie, J.A.; Arya, S.K.; Shaw, G.

J. Virol. 64, 890-901, 1990

A:Title: Molecular characterization of an attenuated human immunodeficiency virus ty

A:Reference number: A33943; MUID:90112662; PMID:2296086

A:Accession: G33943

A:Molecule type: genomic RNA

A:Residues: 1-107 <KUM>

A:Cross-references: GB:M31113; NID:91339798; PIDN:AAB01357.1; PID:9325754

C:Genetics:

A:Gene: rev; trs; art

A:Introns: 24/1

C:Superfamily: AIDS trans-regulatory splicing protein

C:Keywords: splicing protein; transcription regulation

Query Match 10.3%; Score 59.5; DB 1; Length 107;

Best Local Similarity 31.4%; Pred. No. 1.2e+02;

Matches 16; Conservative 5; Mismatches 23; Indels 7; Gaps 1;

QY 41 RGEAPRGVMAVPLLRMEGAPAGPGGRTAACFSCTSRCLSRRPKRSW 91

Db 4 RAEELRRKKLRLRLHQTNPYPOGPG-----TASQRRNRNRWKQRM 47

## RESULT 12

CGRB1S

collagen alpha 1(I) chain - rabbit (tentative sequence) (fragment)

C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 12-Aug-1981 #sequence\_revision 12-Aug-1981 #text\_change 31-Mar-2000

C:Accession: A02856

R:Bornstein, P.; Nesse, R.

Arch. Biochem. Biophys. 138, 443-450, 1970

A:Title: The comparative biochemistry of collagen: the structure of rabbit skin colla

A:Reference number: A02856; MUID:70252720; PMID:4194291

A:Accession: A02856

A:Molecule type: protein

A:Residues: 1-53 <BOR>

A:Experimental source: skin

A:Note: the compositions of CNBr1 and CNBr2 were determined

A:Note: we have positioned residues by homology with rat tendon alpha 1(I) chain

A:Note: it is assumed that the epsilon carbon of Lys-7 is converted to an aldehyde gr

A:Note: the six prolines at the third position of the tripeptide repeating unit (G-X-

C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homolo



```
RESULT 3
S25773
testis-specific protein Mst84Db - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 20-Aug-1999
C:Accession: S25773; B56565
R:Kuhn, R.; Boersch, D.; Glaetzer, K.H.; Schaefer, U.; Schaefer, M.
Mech. Dev. 35, 143-151, 1991
A:Title: A cluster of four genes selectively expressed in the male germ line of Drosophila
A:Reference number: A56565; MUID:92102953; PMID:1684716
A:Accession: S25773
A:Molecule type: DNA
A:Residues: 1-74 <KUH>
A:CROSS-references: EMBL:X67703; NID:gl1072; PIDN:CAA47938.1; PID:gl1074
A:Note: sequence extracted from NCBI backbone (NCBIN:74217, NCBI:P:74220)
C:Genetics:
A:Gene: Mst84Db
A:CROSS-references: FlyBase:FBgn0004173
A:Map position: 3
A:Superfamily: fruit fly testis-specific protein
C:Keywords: spermatogenesis; tandem repeat

Query Match 11.5%; Score 66.5; DB 2; Length 74;
Best Local Similarity 42.1%; Pred. No. 20;
Matches 16; Conservative 0; Mismatches 21; Indels 1; Gaps 1;

QY 62 GAPAGPGGRTAACFCTSRCLS-RRPWKRSWSAGSCPG 98
Db 16 GPGCGPGCGCGCCSCPGCCAPWGCGCGCGCG 53

RESULT 4
A38947
oncofetal-laminin binding collagen - human (fragments)
C:Species: Homo sapiens (man)
C:Date: 13-Jun-1995 #sequence_revision 14-Jul-1995 #text_change 17-Mar-1999
C:Accession: A38947
R:Minatra, I.P.; Andriolo, M.; Basirico, L.; Aquino, A.; Minafra, S.; Boutillon, M.M.; V
Biochem. Biophys. Res. Commun. 207, 852-859, 1995
A:Title: Onco-fetal/laminin-binding collagen from colon carcinoma: Detection of new sequ
A:Reference number: PC2379; MUID:95169134; PMID:7864881
A:Accession: A38947
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-92 <MIN>
A:Experimental source: colorectal carcinomas
C:Superfamily: unassigned collagens

Query Match 11.5%; Score 66.5; DB 2; Length 92;
Best Local Similarity 28.0%; Pred. No. 24;
Matches 23; Conservative 4; Mismatches 28; Indels 27; Gaps 3;

QY 26 PRAAEVPGAGQGGPRGEEAPRGVMAVPLLRMEGAPAGPGGRTAACFCTSRCLSR 85
Db 7 PGAGPGGARGPGPGPLGPPKGG-----WPLGPMGPA----- 40

QY 86 PWRRSWSAGSCGMPHLSPDQG 107
Db 41 PGKGEPCAPGAPGAPG-APGKG 61

RESULT 5
I57012
alpha 2(XI) collagen - mouse (fragment)
C:Species: Mus sp. (mouse)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 13-Aug-1999
C:Accession: I57012
R:Stubbs, L.; Lui, V.C.; Ng, L.J.; Cheah, K.S.
Mamm. Genome 4, 95-103, 1993
A:Title: the alpha 2(XI) collagen gene lies within 8 kb of Pb in the proximal portion of
A:Reference number: I57012; MUID:93160635; PMID:8431643
```

```
A:Accession: I57012
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-99 <RES>
A:CROSS-references: GB:S54563; NID:g265472; PIDN:AAB25343.1; PID:g265473
C:Genetics:
C:Superfamily: collagen alpha 1(V) chain; fibrillar collagen carboxyl-terminal homolo

Query Match 11.0%; Score 63.5; DB 2; Length 99;
Best Local Similarity 37.3%; Pred. No. 48;
Matches 19; Conservative 3; Mismatches 22; Indels 7; Gaps 2;

QY 22 EREVPRAAEVPGAQGGQGGPRG--REEPAGVMAVPLLRMEGAPAGPGGR 70
Db 14 ERGLPGTAGGGLKNEGPGAGPGGAGGAG-----SGPGTGPGR 59

RESULT 6
F72653
hypothetical protein APE0658 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:Accession: F72653
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Ta
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: F72653
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-108 <RAW>
A:CROSS-references: DDBJ:AP000060; NID:g5104188; PIDN:BAA79630.1; PID:d1043416; PID:g
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE0658

Query Match 10.8%; Score 62.5; DB 2; Length 108;
Best Local Similarity 34.2%; Pred. No. 64;
Matches 26; Conservative 7; Mismatches 36; Indels 7; Gaps 4;

QY 37 QQPGRGEEAPRGVMAVPLLRMEGAPAGPGGRTAACFCTSRCLSRPWRKRSWA- 93
Db 20 QRGPGGVFGACGPGPQLQSRSSQAPGGRLYAGCGGIEPRSGAPRRGPGQSTWAAG 79

QY 94 -GSCGMPHLSPDQGR 108
Db 80 VGSHPGS---RPAPGR 92

RESULT 7
B44982
collagen COL4 - pig roundworm
C:Species: Ascaris suum (pig roundworm)
C:Date: 14-May-1993 #sequence_revision 30-Sep-1993 #text_change 20-Sep-1999
C:Accession: B44982
R:Kingston, I.B.; Wainwright, S.M.; Cooper, D.
Mol. Biochem. Parasitol. 37, 137-146, 1989
A:Title: Comparison of collagen gene sequences in Ascaris suum and Caenorhabditis ele
A:Reference number: A44982; MUID:90136710; PMID:2482444
A:Accession: B44982
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-96 <KIN>
A:CROSS-references: GB:J04655; NID:gl59654; PIDN:AAA29370.1; PID:gl59655
C:Superfamily: unassigned collagens

Query Match 10.5%; Score 61; DB 2; Length 96;
Best Local Similarity 27.4%; Pred. No. 79;
Matches 20; Conservative 1; Mismatches 20; Indels 32; Gaps 3;

QY 32 PGAQGGQGGPRGEEAPRGVMAVPLLRMEGAPAGPGGRTAACFCTSRCLSRPWRKRSW 91
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 13:59:09 ; Search time 39 seconds  
(without alignments)  
288.779 Million cell updates/sec

Title: US-09-807-512-2

Perfect score: 579

Sequence: 1 MLMQAELAFMLAQGAMLA.....SWSAGSCPGMPLSPDQGRF 109

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 44098

Minimum DB seq length: 0

Maximum DB seq length: 109

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_76:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	71	12.3	73	JS0514	hypothetical 7.5K
2	70.5	12.2	95	JQ0134	hypothetical 10.8K
3	66.5	11.5	74	S25773	testis-specific pr
4	66.5	11.5	92	A38947	oncofetal-laminin
5	63.5	11.0	99	I57012	alpha 2(XI) collag
6	62.5	10.8	108	F72653	hypothetical prote
7	61	10.5	96	B41982	collagen COL4A - p
8	60.5	10.4	93	I45876	collagen alpha 1(I
9	60.5	10.4	107	B61396	collagen alpha 1(I
10	59.5	10.3	75	PQ0805	core protein - hep
11	59.5	10.3	107	VKLJST	trans-regulatory s
12	58.5	10.1	53	CGRB1S	collagen alpha 1(I
13	58.5	10.1	82	PQ0804	core protein - hep
14	58.5	10.1	86	S00802	collagen alpha 1(X
15	58.5	10.1	103	JQ2283	negatively phytoch
16	58.5	10.1	107	G72715	hypothetical prote
17	58.5	10.1	108	VKLJS2	trans-regulatory s
18	57	9.8	53	I60384	gene T1 protein -
19	57	9.8	78	B44479	collagen alpha 1(X
20	57	9.8	106	B72490	hypothetical prote
21	56.5	9.8	83	I73095	MHC class II beta
22	56	9.7	81	S27366	metallothionein E(
23	56	9.7	100	VKLJG2	trans-regulatory s
24	56	9.7	103	C34770	ORF3 protein - sal
25	55.5	9.6	89	S03805	hypothetical prote
26	55	9.5	58	A45526	collagen homolog A
27	55	9.5	70	T03650	probable malate de
28	55	9.5	76	T09262	glycine-rich cell
29	55	9.5	109	F72724	hypothetical prote

30 54.5 9.4 78 2 JC1406 pigment-dispersing  
31 54.5 9.4 101 2 T09149 late-embryogenesis  
32 54 9.3 102 2 A72507 hypothetical prote  
33 54 9.3 106 2 AC2738 hypothetical prote  
34 53.5 9.2 35 2 B24450 collagen alpha 1(V  
35 53.5 9.2 79 2 B41793 dipeptidyl aminope  
36 53.5 9.2 102 2 A49361 peripherai benzodi  
37 53.5 9.2 105 2 C72580 hypothetical prote  
38 53.5 9.2 107 1 NVWH2F vasopressin / neur  
39 53 9.2 93 2 C75396 hypothetical prote  
40 53 9.2 103 2 H72517 hypothetical prote  
41 53 9.2 106 2 A75263 MHC class II beta  
42 52.5 9.1 82 2 I73096  
43 52.5 9.1 108 2 S41348 genome polypeptin  
44 52.5 9.1 109 2 E70730 hypothetical prote  
45 52 9.0 81 2 S27369 metallothionein E(

#### ALIGNMENTS

##### RESULT 1

JS0514

hypothetical 7.5K protein (TPI intron 1) - chimpanzee

C:Species: Pan troglodytes (chimpanzee)

C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 21-Jul-2000

C:Accession: JS0514

R:Craig, L.C.; Pirtle, I.L.; Gracy, R.W.; Pirtle, R.M.

Gene 99, 217-227, 1991

A:Title: Characterization of the transcription unit and two processed pseudogenes of

A:Reference number: JH0375; MUID:91216462; PMID:2022334

A:Accession: JS0514

A:Molecule type: DNA

A:Residues: 1-73 <CRA>

A:Cross-references: GB:M57946; NID:g176959; PIDN:AAA35439.1; PID:g176961

Query Match

Best Local Similarity 12.3%; Score 71; DB 2; Length 73;

Matches 21; Conservative 2; Mismatches 18; Indels 12; Gaps 2;

QY 33 GAQGGQGRGEEAPRGVMAVPLLRMEGAPAGPGRTAACFSCTSRCLSR 85

DB 32 GARGSGDRGRUTAAAGV-----GEEAGRGLPATC-PLASRAVRR 72

##### RESULT 2

JQ0134

hypothetical 10.8K protein - Pseudomonas aeruginosa

C:Species: Pseudomonas aeruginosa

C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 07-Jun-1996

C:Accession: JQ0134

R:Kato, J.; Chu, L.; Kitano, K.; DeVault, J.D.; Kimbara, K.; Chakrabarty, A.M.; Misra

Gene 84, 31-38, 1989

A:Title: Nucleotide sequence of a regulatory region controlling alginate synthesis in

A:Reference number: JQ0132; MUID:90108714; PMID:2514124

A:Accession: JQ0134

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-95 <KAT>

C:Genetics:

A:Start codon: GTG

Query Match 12.2%; Score 70.5; DB 2; Length 95;

Best Local Similarity 34.1%; Pred. No. 11;

Matches 30; Conservative 4; Mismatches 33; Indels 21; Gaps 7;

QY 23 RRVP-RAAEVPGAQGGQGRGEEAPRGVMAVPLLRMEGAPAG-PGGR-TAACFSCTS 79

DB 8 RRCPRVYAGPG-QNRRAPLRRRPVPPGSR-AGDCRRRRNPPPGDGYRPPGACARCAS 65

QY 80 -----RCLSRPW--KRSW 91

DB 66 RRATGSRRHPRAGRATARTSWTCQRHW 93

```
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 12.4%; Score 72; DB 21; Length 94;
Best Local Similarity 31.9%; Pred No 6.1;
Matches 29; Conservative 13; Mismatches 13; Indels 12; Gaps 5;

Qy 6 EALAFMAQGAMLAQERRYPRAAEVFGAQOQGPGRGEEAPRGVMAVPLLRMEGAP- 64
Db 13 QPIAFPPNPRPMAPRRSPRR- - - - - SRRRSPPRRPPG- - - - - RSP- - - - - RRRGSPL 62

Qy 65 AGPGRTAACFCTSRCLSRPWRKRSWSAGS 95
Db 63 ASPPARRAAS-VARRRARRWRTRSTSS 92
```

Search completed: October 7, 2003, 14:00:25  
Job time : 85 secs

PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140595.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142820.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 22-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157865.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158569.

```

FT Misc-difference 100
FT /label= Unknown
FT /note= "Encoded by NTT"
FT Misc-difference 105
FT /label= Unknown
FT /note= "Encoded by NAA"
FT
XX WO200012732-A2.
XX
XX 09-MAR-2000.
XX
XX 25-AUG-1999; 99WO-US18955.
XX
XX 28-AUG-1998; 98US-0098225.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Bensen RJ;
XX
XX WPI: 2000-256647/22.
XX N-PSDB: AAZ51838.
XX
XX New organelle targeting nucleic acid and amino acid sequences are
XX useful in genetic engineering for modulating the subcellular
XX localization of heterologous proteins in plants -
XX
XX Claim 16; Page 48; 50pp; English.
XX
XX The present sequence is a maize transit peptide which has homology to
XX choporphyrogen oxidase.
XX This sequence can be used in genetic engineering for modulating
XX the subcellular localisation of heterologous proteins in plants. In
XX particular, the transit peptide finds use in the localisation of
XX proteins to plastids and compartments thereof.
XX
XX Sequence 107 AA;
XX
XX Query Match 13.18; Score 76; DB 21; Length 107;
XX Best Local Similarity 33.3%; Pred. No. 2.7;
XX Matches 25; Conservative 7; Mismatches 31; Indels 12; Gaps 4;
XX
XX 33 GAQGQGGPRGEEAPRGVMAVPLLRMEGAPAG--PPGRTAACFCSTSRCLSRPWKRS 90
XX | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
XX 3 GCRGSAPRSRQ-----MAV---GRASSTSGRAAAAAAGSCKTAASSRRPGSTS 53
XX
XX 91 -WSAGSCPGMPHLSP 104
XX | | | | | : |
XX 54 PSTGSCQPTPNAPP 68
XX
XX RESULT 14
XX ID AAM06890 standard; Protein; 94 AA.
XX AC AAM06890;
XX
XX 05-OCT-2001 (first entry)
XX
XX Human foetal protein, SEQ ID NO: 1098.
XX
XX Human; foetal protein; cytostatic; immunosuppressive; immunostimulant;
XX nootropic; neuroprotective; thrombolytic; osteoporotic; antiinflammatory;
XX gene therapy; antisense therapy; cancer; immune disorder;
XX growth disorder; osteoporosis; thrombolytic disorder;
XX nervous system disorder; inflammation.
XX
XX Homo sapiens.
XX
XX WO200155339-A2.
XX
XX 02-AUG-2001.
XX
XX 25-JAN-2001; 2001WO-US02723.

```

```

XX 25-JAN-2000; 2000US-0491404.
XX 15-SEP-2000; 2000US-0663870.
XX 06-NOV-2000; 2000US-0707351.
XX (HYSE-) HYSEQ INC.
XX
XX Yeung G, Ford JE, Boyle BJ, Arterburn MC, Drmanac RA, Tang YT;
XX Liu C, Asundi V, Zhou P, Werhman T;
XX
XX WPI: 2001-465571/50.
XX N-PSDB: AAH94565.
XX
XX Novel fetal proteins useful for the treatment and diagnosis of diseases
XX associated with dysfunction of the protein e.g. cancers, immune
XX disorders, growth disorders, thrombolytic disorders, nervous system
XX disorders and inflammation -
XX
XX Example 4; Page 614; 715pp; English.
XX
XX The invention relates to novel foetal polypeptides encoded by
XX polynucleotides comprising one of 477 sequences fully defined in the
XX specification. The foetal polynucleotides and polypeptides are
XX useful in the treatment and diagnosis of diseases such as cancers,
XX immune disorders, growth disorders (e.g. osteoporosis), thrombolytic
XX disorders, nervous system disorders and inflammation. The present
XX sequence is a polypeptide encoded by a cDNA assembled using
XX an expressed sequence tag (EST) found to be expressed in human
XX foetal tissue cDNA libraries.
XX
XX Sequence 94 AA;
XX
XX Query Match 12.7%; Score 73.5; DB 22; Length 94;
XX Best Local Similarity 32.5%; Pred. No. 4.2;
XX Matches 27; Conservative 3; Mismatches 36; Indels 17; Gaps 4;
XX
XX 35 QGQGGPRGEEA-PRGVMAVPLLRMEGAPAGP-----GRTAACFCSTSRCLSRRP 86
XX | : | | | | : | : | : | : | : | : | : | : | : | : | : | : |
XX 4 EGKLAPRGAERGDPGLWCRPPSLTQVGATHSPGTEGQLLHGLAAA-----LGPCLPSTP 59
XX
XX 87 WK-----RSWSAGSCPGMPHLSP 104
XX | | | | | : | : | : | : | : | : | : | : | : | : | : | : |
XX 60 VKGVCSVRSMKLGVPVCFQVNSP 82
XX
XX RESULT 15
XX AAG18963
XX ID AAG18963 standard; Protein; 94 AA.
XX AC AAG18963;
XX
XX 17-OCT-2000 (first entry)
XX
XX Zea mays protein fragment SEQ ID NO: 20581.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence; corn.
XX
XX Zea mays subsp. mays.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126284.
XX 29-MAR-1999; 99US-0126785.

```



DT 31-JUL-2000 (first entry)  
 XX Synthetic peptide, A5 from CAMEL protein.  
 XX  
 KW CAMEL; CTL-recognised Antigen on MELanoma; cytotoxic T lymphocyte; CTL;  
 KW tumour-associated antigen; LAGE-1; NY-ESO-1; anticancer; melanoma;  
 KW cancer; immunotherapy; immunogenic peptide; immune response; antibody.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 PN WO200023584-A1.  
 XX  
 PD 27-APR-2000.  
 XX  
 XX 15-OCT-1999; 99WO-EP07832.  
 XX  
 XX 16-OCT-1998; 98EP-0119583.  
 XX  
 XX (BOEH ) BOEHRINGER INGELHEIM INT GMBH.  
 PA (UYHO-) UNIV HOSPITAL LEIDEN.  
 XX  
 XX Schrier PI, Aarnoudse CA, Heider K, Klade C;  
 XX  
 XX WPI; 2000-339685/29.  
 XX  
 XX Tumor-associated antigen useful for cancer immunotherapy is encoded by  
 XX the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA -  
 XX Disclosure; Page 17; 73pp; English.  
 XX  
 CC The present sequence is the synthetic peptide A5, derived from the  
 CC hydrophobic regions of human CAMEL (Cytotoxic T lymphocytes (CTL)-  
 CC recognised Antigen on MELanoma), a tumour-associated antigen. It  
 CC corresponds to residues 36-52 of the CAMEL protein and is used to raise  
 CC antibodies. The Lys residue at both termini enable the peptides to be  
 CC linked to the carrier keyhole limpet haemocyanin (KLH) for immunisation.  
 CC CAMEL protein is encoded by the LAGE-1 gene, a tumour-specific antigen.  
 CC It shows strong homology with NY-ESO-1, a melanoma specific tumour  
 CC antigen. The tumour-associated antigen displayed on melanoma cells is  
 CC recognised by cytotoxic T lymphocytes. CAMEL is expressed in tumour cell  
 CC lines, tumour tissues (e.g. breast and lung) and in restricted number of  
 CC healthy tissues. This sequence has anticancer activity. CAMEL tumour  
 CC antigen and immunogenic peptides derived from it are useful for cancer  
 CC immunotherapy. They have the potential to induce an immune response, by  
 CC eliciting a CTL response. The DNA molecule is used to construct  
 CC recombinant or fusion proteins.  
 XX  
 XX Sequence 19 AA;  
 SQ  
 Query Match 15.9%; Score 92; DB 21; Length 19;  
 Best Local Similarity 94.4%; Pred. No. 0.01;  
 Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 35 QGQGGPRGREGAPRGVRM 52  
 Db 1 KGQGGPRGREGAPRGVRM 18  
 :|||||  
 RESULT 10  
 ID AAY70906  
 XX AAY70906 standard; peptide; 19 AA.  
 AC AAY70906;  
 XX  
 XX 31-JUL-2000 (first entry)  
 DT  
 XX Synthetic peptide, B5 from CAMEL protein.  
 DE  
 XX CAMEL; CTL-recognised Antigen on MELanoma; cytotoxic T lymphocyte; CTL;  
 KW tumour-associated antigen; LAGE-1; NY-ESO-1; anticancer; melanoma;  
 KW cancer; immunotherapy; immunogenic peptide; immune response; antibody.  
 XX

OS Homo sapiens.  
 OS Synthetic.  
 PN WO200023584-A1.  
 XX  
 PD 27-APR-2000.  
 XX  
 XX 15-OCT-1999; 99WO-EP07832.  
 XX  
 XX 16-OCT-1998; 98EP-0119583.  
 XX  
 XX (BOEH ) BOEHRINGER INGELHEIM INT GMBH.  
 PA (UYHO-) UNIV HOSPITAL LEIDEN.  
 XX  
 XX Schrier PI, Aarnoudse CA, Heider K, Klade C;  
 XX  
 XX WPI; 2000-339685/29.  
 XX  
 XX Tumor-associated antigen useful for cancer immunotherapy is encoded by  
 XX the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA -  
 XX Disclosure; Page 17; 73pp; English.  
 XX  
 CC The present sequence is the synthetic peptide B5, derived from the  
 CC hydrophobic regions of human CAMEL (Cytotoxic T lymphocytes (CTL)-  
 CC recognised Antigen on MELanoma), a tumour-associated antigen. It  
 CC corresponds to residues 58-73 of the CAMEL protein and is used to raise  
 CC antibodies. The Lys residue at both termini enable the peptides to be  
 CC linked to the carrier keyhole limpet haemocyanin (KLH) for immunisation.  
 CC CAMEL protein is encoded by the LAGE-1 gene, a tumour-specific antigen.  
 CC It shows strong homology with NY-ESO-1, a melanoma specific tumour  
 CC antigen. The tumour-associated antigen displayed on melanoma cells is  
 CC recognised by cytotoxic T lymphocytes. CAMEL is expressed in tumour cell  
 CC lines, tumour tissues (e.g. breast and lung) and in restricted number of  
 CC healthy tissues. This sequence has anticancer activity. CAMEL tumour  
 CC antigen and immunogenic peptides derived from it are useful for cancer  
 CC immunotherapy. They have the potential to induce an immune response, by  
 CC eliciting a CTL response. The DNA molecule is used to construct  
 CC recombinant or fusion proteins.  
 XX  
 XX Sequence 19 AA;  
 SQ  
 Query Match 14.5%; Score 84; DB 21; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 0.067;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 58 RRMEGAPAGPGGRTAA 73  
 Db 3 RRMEGAPAGPGGRTAA 18  
 :|||||  
 RESULT 11  
 ID AAY70904  
 XX AAY70904 standard; peptide; 19 AA.  
 AC AAY70904;  
 XX  
 XX 31-JUL-2000 (first entry)  
 DT  
 XX Synthetic peptide, F4 from CAMEL protein.  
 DE  
 XX CAMEL; CTL-recognised Antigen on MELanoma; cytotoxic T lymphocyte; CTL;  
 KW tumour-associated antigen; LAGE-1; NY-ESO-1; anticancer; melanoma;  
 KW cancer; immunotherapy; immunogenic peptide; immune response; antibody.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 PN WO200023584-A1.  
 XX  
 PD 27-APR-2000.  
 XX  
 XX 15-OCT-1999; 99WO-EP07832.  
 XX





KW viral infection; human immunodeficiency virus; melanoma;  
 KW bacterial infection; Salmonella; Legionella; parasitic infection;  
 KW Trypanosoma; Toxoplasma; Giardia.  
 XX Homo sapiens.  
 OS  
 XX WO200190197-A1.  
 PN  
 XX  
 PD 29-NOV-2001.  
 XX  
 XX 25-MAY-2001; 2001WO-AU00622.  
 PF  
 XX  
 PR 26-MAY-2000; 2000AU-0007761.  
 XX  
 XX (AUSU ) UNIV AUSTRALIAN NAT.  
 PA  
 XX Thomson SA, Ramshaw IA;  
 XX  
 PI WPI; 2002-147575/19.  
 XX  
 DR N-PSDB; ABK36935.  
 XX

New synthetic polypeptides having several different segments of at least one parent polypeptide linked together differently compared to the linkage in the parent polypeptide, for inducing immune response against a pathogen or cancer

Example 3; Fig 27; 364pp; English.

The invention relates to a new synthetic polypeptide (I) comprising several different segments of at least one parent polypeptide linked together in a different relationship relative to their linkage in the parent polypeptide to impede, abrogate or otherwise alter at least one function associated with the parent polypeptide and for inducing an immune response against a pathogen or cancer. Also included are a synthetic polynucleotide encoding and a computer system for designing the synthetic polypeptides. The synthetic polypeptides and polynucleotides are referred to as a Savine. The synthetic polypeptide is useful for modulating immune responses preferably directed against a pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone liver, oesophagus, brain, testicle, uterus), as potentiating agents. Compositions comprising the polypeptide may be used in the treatment or prophylaxis against viral (such as infections caused by HIV (human immunodeficiency virus), hepatitis, influenza, Japanese encephalitis virus, Epstein-Barr virus and respiratory syncytial virus), bacterial (e.g., infections caused by Neisseria, Meningococcal, Haemophilus, Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic (e.g., infections caused by Plasmodium, Schistosoma, Leishmania, Trypanosoma, Toxoplasma and Giardia) infections. The present sequence is a peptide derived from a parent protein used to construct a savine of the invention..

XX Sequence 30 AA;

Query Match 26.3%; Score 152; DB 23; Length 30;  
 Best Local Similarity 100.0%; Pred. No. 1e-08;  
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 QGAMLAQERRVPRAAEVPGAQGGQGRGR 43  
 |||||  
 Db 1 QGAMLAQERRVPRAAEVPGAQGGQGRGR 30

RESULT 6  
 AAU85116  
 ID AAU85116 standard; Peptide; 30 AA.  
 XX  
 AC AAU85116;

DT 08-MAY-2002 (first entry)  
 XX  
 DE Human NYN501b segment 3.  
 XX

KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;  
 KW viral infection; human immunodeficiency virus; melanoma;  
 KW bacterial infection; Salmonella; Legionella; parasitic infection;  
 KW Trypanosoma; Toxoplasma; Giardia.  
 XX Homo sapiens.  
 OS  
 XX WO200190197-A1.  
 PN  
 XX  
 PD 29-NOV-2001.  
 XX  
 XX 25-MAY-2001; 2001WO-AU00622.  
 PF  
 XX  
 PR 26-MAY-2000; 2000AU-0007761.  
 XX  
 XX (AUSU ) UNIV AUSTRALIAN NAT.  
 PA  
 XX Thomson SA, Ramshaw IA;  
 XX  
 PI WPI; 2002-147575/19.  
 XX  
 DR N-PSDB; ABK36936.  
 XX

New synthetic polypeptides having several different segments of at least one parent polypeptide linked together differently compared to the linkage in the parent polypeptide, for inducing immune response against a pathogen or cancer

Example 3; Fig 27; 364pp; English.

The invention relates to a new synthetic polypeptide (I) comprising several different segments of at least one parent polypeptide linked together in a different relationship relative to their linkage in the parent polypeptide to impede, abrogate or otherwise alter at least one function associated with the parent polypeptide and for inducing an immune response against a pathogen or cancer. Also included are a synthetic polynucleotide encoding and a computer system for designing the synthetic polypeptides. The synthetic polypeptides and polynucleotides are referred to as a Savine. The synthetic polypeptide is useful for modulating immune responses preferably directed against a pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone liver, oesophagus, brain, testicle, uterus), as potentiating agents. Compositions comprising the polypeptide may be used in the treatment or prophylaxis against viral (such as infections caused by HIV (human immunodeficiency virus), hepatitis, influenza, Japanese encephalitis virus, Epstein-Barr virus and respiratory syncytial virus), bacterial (e.g., infections caused by Neisseria, Meningococcal, Haemophilus, Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic (e.g., infections caused by Plasmodium, Schistosoma, Leishmania, Trypanosoma, Toxoplasma and Giardia) infections. The present sequence is a peptide derived from a parent protein used to construct a savine of the invention.

XX Sequence 30 AA;

Query Match 23.0%; Score 133; DB 23; Length 30;  
 Best Local Similarity 79.4%; Pred. No. 9.3e-07;  
 Matches 27; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

QY 29 AEVPGAGGQGGPRGEEAPRGVPMVPLLRMEG 62  
 |||||  
 Db 1 AEVPGAGGQGGPRGEEAPRGVPMVPLLRMEG 30

RESULT 7  
 AAU85114  
 ID AAU85114 standard; Peptide; 30 AA.  
 XX  
 AC AAU85114;

DT 08-MAY-2002 (first entry)  
 XX  
 DE Human NYN501b segment 1.  
 XX

```
XX NY-ESO-1; CAMEL; CTL-recognised Antigen on MELanoma; human; cancer; CTL;
KW cytotoxic T lymphocyte; tumour-associated antigen; LAGE-1; anticancer;
KW melanoma; immunotherapy; immune response; variant.
XX
OS Homo sapiens.
XX WO200023584-A1.
XX
XX 27-APR-2000.
XX
XX 15-OCT-1999; 99WO-EP07832.
XX
XX 16-OCT-1998; 98EP-0119583.
XX
XX (BOEH) BOEHRINGER INGELHEIM INT GMBH.
PA (UYHO-) UNIV HOSPITAL LEIDEN.
XX
XX Schrier PI, Aarnoudse CA, Heider K, Klade C;
XX
XX WPI: 2000-339685/29.
XX N-PSDB; AAD00152.
XX
XX Tumor-associated antigen useful for cancer immunotherapy is encoded by
XX the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA -
XX
XX Example 3; Page 65; 73pp; English.
XX
XX The present sequence is the human NY-ESO-1 short variant protein, a
XX tumour antigen, identified by screening an esophagus carcinoma cDNA
XX library. This protein is derived from open reading frame (ORF)-2,
XX that differs from the CAMEL protein (Cytotoxic T lymphocytes (CTL)-
XX recognised Antigen on MELanoma), a tumour-associated antigen, by only its
XX last 5 amino acids. It contains epitopes of tumour specific T-cells.
XX NY-ESO-1 is expressed in different tumour types, but not in healthy
XX tissues except in testis. The tumour-associated antigen displayed on
XX melanoma cells is recognised by cytotoxic T lymphocytes. This sequence
XX has anticancer activity. CAMEL tumour antigen and immunogenic peptides
XX derived from it are useful for cancer immunotherapy. They have the
XX potential to induce an immune response, by eliciting a CTL response.
XX The DNA molecule is used for the construction of recombinant or fusion
XX proteins.
XX
XX Sequence 58 AA;
XX
XX Query Match 45.6%; Score 264; DB 21; Length 58;
XX Best Local Similarity 88.7%; Pred. No. 5.5e-20;
XX Matches 55; Conservative 2; Mismatches 1; Indels 4; Gaps 1;
XX
XX QY 1 MLMAQALAFMLAQGAMLAQAQRRVPRAAEVPGAQGGQGGPRGEEAPRGVMAVPLLRM 60
XX Db 1 MLMAQALAFMLAQGAMLAQAQRRVPRAAEVPGAQGGQGGPRGEEAPRGVMAA----RL 56
XX
XX QY 61 EG 62
XX Db :|
XX 57 QG 58
XX
XX RESULT 4
XX AAU84819
XX
XX AAU84819 standard; Protein: 58 AA.
XX
XX AAU84819;
XX
XX 08-MAY-2002 (first entry)
XX
XX Human NYN501b consensus sequence.
XX
XX Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;
XX viral infection; human immunodeficiency virus; melanoma;
XX bacterial infection; Salmonella; Legionella; parasitic infection;
XX Trypanosoma; Toxoplasma; Giardia.
XX
XX
```

```
OS Homo sapiens.
XX WO200190197-A1.
XX
XX 29-NOV-2001.
XX
XX 25-MAY-2001; 2001WO-AU00622.
XX
XX 26-MAY-2000; 2000AU-0007761.
XX
XX (AUSU) UNIV AUSTRALIAN NAT.
XX
XX Thomson SA, Ramshaw IA;
XX WPI: 2002-147575/19.
XX
XX New synthetic polypeptides having several different segments of at
XX least one parent polypeptide linked together differently compared to
XX the linkage in the parent polypeptide, for inducing immune response
XX against a pathogen or cancer -
XX
XX Example 3; Fig 27; 364pp; English.
XX
XX The invention relates to a new synthetic polypeptide (I) comprising
XX several different segments of at least one parent polypeptide linked
XX together in a different relationship relative to their linkage in the
XX parent polypeptide to impede, abrogate or otherwise alter at least one
XX function associated with the parent polypeptide and for inducing an
XX immune response against a pathogen or cancer. Also included are a
XX synthetic polynucleotide encoding and a computer system for
XX designing the synthetic polypeptides. The synthetic polypeptide is
XX polynucleotides are referred to as a Savine. The synthetic polypeptide is
XX useful for modulating immune responses preferably directed against a
XX pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, bone
XX colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone
XX liver, oesophagus, brain, testicle, uterus), as potentiating agents,
XX Compositions comprising the polypeptide may be used in the treatment or
XX prophylaxis against viral (such as infections caused by HIV (human
XX immunodeficiency virus), hepatitis, influenza, Japanese encephalitis
XX virus, Epstein-Barr virus and respiratory syncytial virus), bacterial
XX (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,
XX Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic
XX (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,
XX Trypanosoma, Toxoplasma and Giardia) infections. The present
XX sequence is a consensus sequence for a parent protein used to design a
XX Savine of the invention.
XX
XX Sequence 58 AA;
XX
XX Query Match 45.6%; Score 264; DB 23; Length 58;
XX Best Local Similarity 88.7%; Pred. No. 5.5e-20;
XX Matches 55; Conservative 2; Mismatches 1; Indels 4; Gaps 1;
XX
XX QY 1 MLMAQALAFMLAQGAMLAQAQRRVPRAAEVPGAQGGQGGPRGEEAPRGVMAVPLLRM 60
XX Db 1 MLMAQALAFMLAQGAMLAQAQRRVPRAAEVPGAQGGQGGPRGEEAPRGVMAA----RL 56
XX
XX QY 61 EG 62
XX Db :|
XX 57 QG 58
XX
XX RESULT 5
XX AAU85115
XX
XX AAU85115 standard; Peptide: 30 AA.
XX
XX AAU85115;
XX
XX 08-MAY-2002 (first entry)
XX
XX Human NYN501b segment 2.
XX
XX Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;
XX
```

FT /label= CAMEL16  
 FT /note= "Specific for HLA-A2"  
 FT 17..25  
 FT /label= CAMEL17  
 FT /note= "Specific for HLA-A2"  
 FT 51..59  
 FT /label= Immunogenic\_peptide  
 FT /note= "Specific for HLA-A3101"  
 FT 101..109  
 FT /label= Immunogenic\_peptide

PN W0200023584-A1.

PD 27-APR-2000.

XX 15-OCT-1999; 99WO-EP07832.

XX 16-OCT-1998; 98EP-0119583.

XX (BOEH ) BOEHRINGER INGELHEIM INT GMBH.  
 XX (UYHO-) UNIV HOSPITAL LEIDEN.

XX Schrier PI, Aarnoudse CA, Heider K, Klade C;

DR WPI; 2000-339685/29.

DR N-PSDB; AAD00149.

XX Tumor-associated antigen useful for cancer immunotherapy is encoded by  
 the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA -

XX Claim 1; Page 55; 73pp; English.

CC The present protein sequence is the human tumour-associated antigen CAMEL  
 CC (cytotoxic T lymphocytes (CTL)-recognised Antigen on MELANOMA). CAMEL  
 CC protein is encoded by the LAGE-1 gene, a tumour-specific antigen. It is  
 CC different from the LAGE-1 protein, since it is translated from a  
 CC different open reading frame (ORF-1). It shows strong homology with  
 CC NY-ESO-1, a melanoma specific tumour antigen. The tumour-associated  
 CC antigen displayed on melanoma cells is recognised by cytotoxic T  
 CC lymphocytes. CAMEL is expressed in tumour cell lines, tumour tissues  
 CC (e.g. breast and lung) and in restricted number of healthy tissues. This  
 CC sequence has anticancer activity. CAMEL tumour antigen and immunogenic  
 CC peptides derived from it are useful for cancer immunotherapy. They have  
 CC the potential to induce an immune response, by eliciting a CTL response.  
 CC The DNA molecule is used to construct recombinant or fusion proteins.

XX Sequence 109 AA;

Query Match 100.0%; Score 579; DB 21; Length 109;  
 Best Local Similarity 100.0%; Pred.No.3.2e-52;  
 Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLMAQALAFMAQAGMALAAQERRVPRAAEPVGAQGGQGGPRGEEAPRGVMAVPLLRM 60

DB 1 MLMAQALAFMAQAGMALAAQERRVPRAAEPVGAQGGQGGPRGEEAPRGVMAVPLLRM 60

QY 61 EGAPAGPGGRTAACFSCTSCLSRPWRKRSWSAGSCPGMPLSPDQGRF 109

DB 61 EGAPAGPGGRTAACFSCTSCLSRPWRKRSWSAGSCPGMPLSPDQGRF 109

RESULT 2

AAAY05966

ID AAY05966 standard; Protein; 58 AA.

XX AAAY05966;

XX 16-AUG-1999 (first entry)

XX Human cancer antigen NY ESO-1/CAG-3 ORF2 protein.

XX NY ESO-1/CAG-3 gene; CAG-3 gene; cancer antigen; human; leukaemia;  
 KW non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;

KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
 KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
 KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
 KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
 KW vaccine; ORF2.

XX Homo sapiens.

OS WO9918206-A2.

PN 15-APR-1999.

PD 21-SEP-1998; 98WO-US19609.

XX 08-OCT-1997; 97US-0061428.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES..

XX Rosenberg SA, Wang RF;

XX WPI; 1999-277270/23.

XX N-PSDB; AAX58599.

XX Cancer antigen NY ESO1/CAG-3

XX Claim 5; Fig 3A; 88pp; English.

XX The present sequence represents the ORF2 protein encoded by

XX open reading frame 2 of the human ESO-1/CAG-3 (or CAG-3) gene.

XX CAG-3 is a new and potent tumour antigen capable of eliciting an

XX antigen specific immune response by T cells. Cancer peptides

XX comprising ORF2, ORF1 (see AAY05965), portions of these peptides and

XX their variants (see AAY05967-87), are useful as cancer vaccines that

XX protect the recipient from development of cancer. The invention

XX provides vectors and host cells (also useful as vaccines); a

XX method of diagnosis of cancer or precancer; a transgenic animal;

XX antisense oligonucleotides that inhibit expression of the cancer

XX peptide or tumour antigen; antibodies reacting with the CAG-3

XX cancer peptide, useful in diagnostic and detection assays; and

XX methods for preventing or inhibiting cancer by administering a

XX cancer peptide, with or without an HLA molecule. The cancer

XX peptides form part of, or are derived from, cancers such as

XX primary or metastatic melanoma, thymoma, lymphoma, sarcoma, lung

XX cancer, liver cancer, leukaemia, uterine cancer, cervical cancer,

XX bladder cancer, kidney cancer and adenocarcinomas such as breast,

XX prostate, ovarian, pancreatic and thyroid cancers. Melanoma is

XX treated by inducing cancer-specific T cells in vitro for subsequent

XX return to a patient.

XX Sequence 58 AA;

QY Query Match 45.6%; Score 264; DB 20; Length 58;

Best Local Similarity 88.7%; Pred.No.5.5e-20;

Matches 55; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

QY 1 MLMAQALAFMAQAGMALAAQERRVPRAAEPVGAQGGQGGPRGEEAPRGVMAVPLLRM 60

DB 1 MLMAQALAFMAQAGMALAAQERRVPRAAEPVGAQGGQGGPRGEEAPRGVMAVPLLRM 60

QY 61 EG 62

DB 57 QG 58

RESULT 3

AAAY0863.

ID AAY0863 standard; Protein; 58 AA.

XX AAAY0863;

XX 31-JUL-2000 (first entry)

XX Human tumour antigen, NY-ESO-1 short variant protein.

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: October 7, 2003, 13:32:48 ; Search time 84 Seconds  
(without alignments)  
205.967 Million cell updates/sec

Title: US-09-807-512-2

Perfect score: 579

Sequence: 1 MLMAQEAFLAQMAGMLAA.....SWSAGSCPCPHLSPDQGRF 109

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 737625

Minimum DB seq length: 0

Maximum DB seq length: 109

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*

ID	Score	Match	Length	DB	ID	Description
1:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*					
2:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*					
3:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*					
4:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*					
5:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*					
6:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*					
7:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*					
8:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*					
9:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*					
10:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*					
11:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*					
12:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*					
13:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*					
14:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*					
15:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*					
16:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*					
17:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*					
18:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*					
19:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*					
20:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*					
21:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*					
22:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*					
23:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*					
24:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*					

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	579	100.0	109	21	AA1980	Human CTL-recog
2	264	45.6	58	20	AA1981	Human cancer anti
3	264	45.6	58	21	AA1982	Human tumour anti
4	264	45.6	58	23	AA1983	Human tumour anti
5	152	26.3	30	23	AA1984	Human tumour anti
6	133	23.0	30	23	AA1985	Human tumour anti
7	131	22.6	30	23	AA1986	Human tumour anti
8	92.5	16.0	101	22	AA1987	Novel human diagno
9	92	15.9	19	21	AA1988	Synthetic peptide,

10	84	14.5	19	21	AA1989	Synthetic peptide,
11	81	14.0	19	21	AA1990	Synthetic peptide,
12	77	13.3	80	21	AA1991	A human collagen 1
13	76	13.1	107	21	AA1992	Maize plastid targ
14	73.5	12.7	94	22	AA1993	Human foetal prote
15	72	12.4	94	21	AA1994	zee mays protein f
16	71	12.3	87	24	AA1995	HCV related recomb
17	70.5	12.2	104	22	AA1996	Novel human diagno
18	70.5	12.2	106	22	AA1997	Novel human diagno
19	70	12.1	73	22	AA1998	Propionibacterium
20	70	12.1	84	22	AA1999	Novel human diagno
21	70	12.1	104	21	AA2000	zee mays protein f
22	70	12.1	104	21	AA2001	Novel human diagno
23	69.5	12.0	109	22	AA2002	Human secreted pro
24	69.5	12.0	104	22	AA2003	Human immune/haema
25	69	11.9	78	23	AA2004	Human peptide enco
26	68.5	11.8	53	22	AA2005	Propionibacterium
27	68.5	11.8	95	24	AA2006	Human FBXO31-4 (P-
28	68.5	11.8	104	22	AA2007	Propionibacterium
29	68	11.7	109	23	AA2008	Human secreted pro
30	67	11.6	108	22	AA2009	Propionibacterium
31	66.5	11.5	81	19	AA2010	Human neurofilamen
32	66.5	11.5	91	23	AA2011	Human ORF1512 prot
33	66.5	11.5	96	22	AA2012	Human immune/haema
34	66.5	11.5	104	22	AA2013	Human polypeptide
35	66.5	11.5	104	24	AA2014	Maize granule boun
36	66.5	11.5	104	24	AA2015	Maize granule boun
37	66.5	11.5	104	24	AA2016	Maize granule boun
38	66.5	11.5	107	21	AA2017	Fragment of human
39	66	11.4	60	14	AA2018	CLP 3.1 monomer e
40	66	11.4	72	22	AA2019	Propionibacterium
41	66	11.4	85	22	AA2020	Human immune/haema
42	66	11.4	92	22	AA2021	Novel human secret
43	65.5	11.3	80	23	AA2022	Human ORF2174 prot
44	65.5	11.3	103	23	AA2023	Human DIRHP polype
45	65	11.2	56	14	AA2024	Collagen-like poly

## ALIGNMENTS

RESULT 1

AA1989

ID AA1989 standard; Protein; 109 AA.

XX AA1989

AC AA1989

31-JUL-2000 (first entry)

Human CTL-recognised Antigen on MELANOMA (CAMEL) protein.

CAMEL; CTL-recognised Antigen on MELANOMA: cytotoxic T lymphocyte; CTL;

tumour-associated antigen; LAGE-1; NY-ESO-1; anticancer; melanoma;

cancer; immunotherapy; immunogenic peptide; immune response.

OS Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

**THIS PAGE BLANK (USPTO)**

---

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:14:20 ; Search time 24 Seconds  
(without alignments)  
13.350 Million cell updates/sec

Title: us-09-807-512-11

Perfect score: 50

Sequence: 1 MLMAQERLAFL 11

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 136899 seqs, 29126274 residues

Total number of hits satisfying chosen parameters: 27083

Minimum DB seq length: 0

Maximum DB seq length: 11

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_RA\_New:\*

- 1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	19	38.0	8	1	PCT-US03-27706-385
2	19	38.0	8	5	US-09-883-859-46
3	19	38.0	8	6	US-10-657-022-385
4	19	38.0	9	1	PCT-US03-27706-384
5	19	38.0	9	5	US-09-872-186-6
6	19	38.0	9	5	US-09-856-988B-29
7	19	38.0	9	6	US-10-647-005-27
8	19	38.0	9	6	US-10-657-022-384
9	19	38.0	9	7	US-06-485-717-87
10	19	38.0	10	1	PCT-US03-27706-383
11	19	38.0	10	1	PCT-US03-27706-386
12	19	38.0	10	1	PCT-US03-27706-387
13	19	38.0	10	5	US-09-239-043D-19
14	19	38.0	10	5	US-09-239-043D-696
15	19	38.0	10	5	US-09-239-043D-1535
16	19	38.0	10	5	US-09-239-043D-1798
17	19	38.0	10	6	US-10-657-022-383
18	19	38.0	10	6	US-10-657-022-386
19	19	38.0	10	6	US-10-657-022-387
20	19	38.0	10	6	US-10-658-180-213
21	19	38.0	10	6	US-10-658-180-217
22	19	38.0	10	6	US-10-363-990-19
23	19	38.0	10	6	US-10-363-990-731
24	19	38.0	10	6	US-10-363-990-1824
25	19	38.0	10	6	US-10-363-990-2125
26	19	38.0	11	5	US-09-239-043D-20

27	19	38.0	11	5	US-09-239-043D-697	Sequence 697, App
28	19	38.0	11	5	US-09-239-043D-1315	Sequence 1315, Ap
29	19	38.0	11	5	US-09-239-043D-1536	Sequence 1536, Ap
30	19	38.0	11	6	US-10-363-990-20	Sequence 20, Appl
31	19	38.0	11	6	US-10-363-990-732	Sequence 732, App
32	19	38.0	11	6	US-10-363-990-1936	Sequence 1936, Ap
33	19	38.0	11	6	US-10-363-990-2126	Sequence 2126, Ap
34	18	36.0	8	5	US-09-390-061D-250	Sequence 250, App
35	18	36.0	8	5	US-09-390-061D-2049	Sequence 2049, Ap
36	18	36.0	9	1	PCT-US03-23503-29	Sequence 29, Appl
37	18	36.0	9	1	PCT-US03-27706-66	Sequence 66, Appl
38	18	36.0	9	5	US-09-390-061D-613	Sequence 613, App
39	18	36.0	9	5	US-09-390-061D-2050	Sequence 2050, Ap
40	18	36.0	9	5	US-09-390-061D-2238	Sequence 2238, Ap
41	18	36.0	9	5	US-09-390-061D-2254	Sequence 2254, Ap
42	18	36.0	9	5	US-09-390-061D-2697	Sequence 2697, Ap
43	18	36.0	9	5	US-09-390-061D-3417	Sequence 3417, Ap
44	18	36.0	9	5	US-09-390-061D-3418	Sequence 3418, Ap
45	18	36.0	9	5	US-09-453-174-14	Sequence 14, Appl

#### ALIGNMENTS

RESULT 1  
PCT-US03-27706-385  
; Sequence 385, Application PC/TUS0327706  
; GENERAL INFORMATION:  
; APPLICANT: MANNKIND CORPORATION  
; APPLICANT: SIMARD, John J. L.  
; APPLICANT: DIAMOND, David C.  
; APPLICANT: LIU, Liping  
; APPLICANT: LIU, Zheng  
; TITLE OF INVENTION: EPITOPE SEQUENCES  
; FILE REFERENCE: MANK.032VPC  
; CURRENT APPLICATION NUMBER: PCT/US03/27706  
; CURRENT FILING DATE: 2003-09-05  
; PRIOR APPLICATION NUMBER: US 60/409123  
; PRIOR FILING DATE: 2002-09-06  
; NUMBER OF SEQ ID NOS: 610  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 385  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Homosapiens  
PCT-US03-27706-385

Query Match 38.0%; Score 19; DB 1; Length 8;  
Best Local Similarity 66.7%; Pred. No. 1.2e+05;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 EALAF 11  
| | | |  
Db 3 ERLAY 8

RESULT 2  
US-09-883-859-46  
; Sequence 46, Application US/09883859  
; GENERAL INFORMATION:  
; APPLICANT: Sicheiri et al.  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURE OF A STERILE ALPHA MOTIF DOMAIN  
; FILE REFERENCE: MTSI-P01-010  
; CURRENT APPLICATION NUMBER: US/09/883,859  
; CURRENT FILING DATE: 2001-06-18  
; PRIOR APPLICATION NUMBER: PCT/CA99/01209  
; PRIOR FILING DATE: 1999-12-17  
; NUMBER OF SEQ ID NOS: 126  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 46  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence

; FEATURE:  
; OTHER INFORMATION: peptide molecule that modulates SAM domain function  
US-09-883-859-46

Query Match 38.0%; Score 19; DB 5; Length 8;  
Best Local Similarity 57.1%; Pred. No. 1.2e+05;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 MAQEALA 9  
| : | : |  
DB 2 MSQDDLA 8

## RESULT 3

US-10-657-022-385  
; Sequence 385, Application US/10657022  
; GENERAL INFORMATION:  
; APPLICANT: Simard, John J. L.  
; APPLICANT: Diamond, David C.  
; APPLICANT: Liu, Liping  
; APPLICANT: Liu, Zheng  
; TITLE OF INVENTION: EPTIPOE SEQUENCES  
; FILE REFERENCE: MANNK.032A  
; CURRENT APPLICATION NUMBER: US/10/657,022  
; CURRENT FILING DATE: 2003-09-04  
; PRIOR APPLICATION NUMBER: 60/409123  
; PRIOR FILING DATE: 2002-09-06  
; NUMBER OF SEQ ID NOS: 610  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 385  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Homosapiens  
US-10-657-022-385

Query Match 38.0%; Score 19; DB 6; Length 8;  
Best Local Similarity 66.7%; Pred. No. 1.2e+05;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 EALAFLL 11  
| : | : |  
DB 3 ERLAYL 8

## RESULT 4

PCT-US03-27706-384  
; Sequence 384, Application PC/TUS0327706  
; GENERAL INFORMATION:  
; APPLICANT: MANNKIND CORPORATION  
; APPLICANT: SIMARD, John J. L.  
; APPLICANT: DIAMOND, David C.  
; APPLICANT: LIU, Liping  
; APPLICANT: LIU, Zheng  
; TITLE OF INVENTION: EPTIPOE SEQUENCES  
; FILE REFERENCE: MANNK.032VPC  
; CURRENT APPLICATION NUMBER: PCT/US03/27706  
; CURRENT FILING DATE: 2003-09-05  
; PRIOR APPLICATION NUMBER: US 60/409123  
; PRIOR FILING DATE: 2002-09-06  
; NUMBER OF SEQ ID NOS: 610  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 384  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homosapiens  
PCT-US03-27706-384

Query Match 38.0%; Score 19; DB 1; Length 9;  
Best Local Similarity 66.7%; Pred. No. 1.2e+05;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 EALAFLL 11  
| : | : |

DB 4 ERLAYL 9

## RESULT 5

US-09-872-186-6  
; Sequence 6, Application US/09872186  
; GENERAL INFORMATION:  
; APPLICANT: John R. Subjeck  
; APPLICANT: Robert A. Henderson  
; APPLICANT: Elizabeth A. Repasky  
; APPLICANT: Latif Kazim  
; APPLICANT: Xiang-Yang Wang  
; TITLE OF INVENTION: STRESS PROTEIN COMPOSITIONS AND METHODS  
; TITLE OF INVENTION: FOR PREVENTION AND TREATMENT OF CANCER AND INFECTIOUS  
; TITLE OF INVENTION: DISEASE  
; FILE REFERENCE: 126.7US11  
; CURRENT APPLICATION NUMBER: US/09/872,186  
; CURRENT FILING DATE: 2001-06-01  
; PRIOR APPLICATION NUMBER: 09/676,340  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/156,821  
; PRIOR FILING DATE: 1999-09-30  
; PRIOR APPLICATION NUMBER: 60/163,168  
; PRIOR FILING DATE: 1999-11-02  
; PRIOR APPLICATION NUMBER: 60/215,497  
; PRIOR FILING DATE: 2000-06-30  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-872-186-6

Query Match 38.0%; Score 19; DB 5; Length 9;  
Best Local Similarity 80.0%; Pred. No. 1.2e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 ALAFL 11  
| : | : |  
DB 5 SLAFL 9

## RESULT 6

US-09-856-988B-29  
; Sequence 29, Application US/09856988B  
; GENERAL INFORMATION:  
; APPLICANT: SCHLOM, JEFFREY  
; APPLICANT: HODGE, JAMES  
; APPLICANT: PANICALI, DENNIS  
; TITLE OF INVENTION: A RECOMBINANT VECTOR EXPRESSING MULTIPLE COSTIMULATORY  
; TITLE OF INVENTION: MOLECULES AND USES THEREOF  
; FILE REFERENCE: 38163-0131  
; CURRENT APPLICATION NUMBER: US/09/856,988B  
; CURRENT FILING DATE: 2001-09-24  
; PRIOR APPLICATION NUMBER: PCT/US99/26866  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR APPLICATION NUMBER: 60/111,582  
; PRIOR FILING DATE: 1998-12-09  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 29  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: peptide  
US-09-856-988B-29

Query Match 38.0%; Score 19; DB 5; Length 9;  
Best Local Similarity 80.0%; Pred. No. 1.2e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;



QY 7 ALAFL 11  
:||||  
Db 5 SLAFL 9

## RESULT 7

US-10-647-005-27  
; Sequence 27, Application US/10647005  
; GENERAL INFORMATION:  
; APPLICANT: Cheever, Martin A.  
; Disis, Mary L.  
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN  
; FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
; HER-2/neu ONCOGENE IS ASSOCIATED  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed IP Law Group PLLC  
; STREET: 701 Fifth Avenue Suite 6300  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: US  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/647,005  
; FILING DATE: 21-Aug-2003  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharkey, Richard G.  
; REGISTRATION NUMBER: 32,629  
; REFERENCE/DOCKET NUMBER: 920010.448C10  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: <Unknown>  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 27:  
US-10-647-005-27

Query Match 38.0%; Score 19; DB 6; Length 9;  
Best Local Similarity 80.0%; Pred. No. 1.2e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 ALAFL 11  
:||||  
Db 5 SLAFL 9

## RESULT 8

US-10-657-022-384  
; Sequence 384, Application US/10657022  
; GENERAL INFORMATION:  
; APPLICANT: Simard, John J. L.  
; APPLICANT: Diamond, David C.  
; APPLICANT: Liu, Liping  
; APPLICANT: Liu, Zheng  
; TITLE OF INVENTION: EPITOPE SEQUENCES  
; FILE REFERENCE: MANNK.032A  
; CURRENT APPLICATION NUMBER: US/10/657,022  
; CURRENT FILING DATE: 2003-09-04  
; PRIOR APPLICATION NUMBER: 60/409123  
; PRIOR FILING DATE: 2002-09-06  
; NUMBER OF SEQ ID NOS: 610  
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 384  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homosapiens  
US-10-657-022-384

Query Match 38.0%; Score 19; DB 6; Length 9;  
Best Local Similarity 66.7%; Pred. No. 1.2e+05;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 EALAFL 11  
:||||  
Db 4 ERLAYL 9

## RESULT 9

US-60-485-717-87  
; Sequence 87, Application US/60485717  
; GENERAL INFORMATION:  
; APPLICANT: Bachmann, Martin F  
; APPLICANT: Schwarz, Katrin  
; TITLE OF INVENTION: SYNERGISTIC ENHANCEMENT OF T CELL RESPONSES BY CPGS AND  
; TOLL-LIKE RECEPTOR 4 LIGANDS OR SAPONIN: METHOD OF PREPARATION  
; FILE REFERENCE: 1700.0430000  
; CURRENT APPLICATION NUMBER: US/60/485,717  
; CURRENT FILING DATE: 2003-07-10  
; NUMBER OF SEQ ID NOS: 132  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 87  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-60-485-717-87

Query Match 38.0%; Score 19; DB 7; Length 9;  
Best Local Similarity 80.0%; Pred. No. 1.2e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 ALAFL 11  
:||||  
Db 5 SLAFL 9

RESULT 10  
PCT-US03-27706-383  
; Sequence 383, Application PC/TUS0327706  
; GENERAL INFORMATION:  
; APPLICANT: MANNKIND CORPORATION  
; APPLICANT: SIMARD, John J. L.  
; APPLICANT: DIAMOND, David C.  
; APPLICANT: LIU, Liping  
; APPLICANT: LIU, Zheng  
; TITLE OF INVENTION: EPITOPE SEQUENCES  
; FILE REFERENCE: MANNK.032VPC  
; CURRENT APPLICATION NUMBER: PCT/US03/27706  
; CURRENT FILING DATE: 2003-09-05  
; PRIOR APPLICATION NUMBER: US 60/409123  
; PRIOR FILING DATE: 2002-09-06  
; NUMBER OF SEQ ID NOS: 610  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 383  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homosapiens  
PCT-US03-27706-383

Query Match 38.0%; Score 19; DB 1; Length 10;  
Best Local Similarity 66.7%; Pred. No. 1.2e+03;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 EALAFL 11  
:||||  
Db 5 ERLAYL 10

## RESULT 11

PCT-US03-27706-386  
; Sequence 386, Application PC/TUS0327706  
; GENERAL INFORMATION:  
; APPLICANT: MANKIND CORPORATION  
; APPLICANT: SIMARD, John J. L.  
; APPLICANT: DIAMOND, David C.  
; APPLICANT: LIU, Zheng  
; APPLICANT: LIU, Zheng  
; TITLE OF INVENTION: EPITOPE SEQUENCES  
; FILE REFERENCE: MANK.032VPC  
; CURRENT APPLICATION NUMBER: PCT/US03/27706  
; CURRENT FILING DATE: 2003-09-05  
; PRIOR APPLICATION NUMBER: US 60/409123  
; PRIOR FILING DATE: 2002-09-06  
; NUMBER OF SEQ ID NOS: 610  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 386  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homosapiens  
PCT-US03-27706-386

Query Match 38.0%; Score 19; DB 1; Length 10;  
Best Local Similarity 66.7%; Pred. No. 1.2e+03;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 6 EALAFI 11  
| | | |  
Db 3 ERLAYL 8

## RESULT 12

PCT-US03-27706-387  
; Sequence 387, Application PC/TUS0327706  
; GENERAL INFORMATION:  
; APPLICANT: MANKIND CORPORATION  
; APPLICANT: SIMARD, John J. L.  
; APPLICANT: DIAMOND, David C.  
; APPLICANT: LIU, Liping  
; APPLICANT: LIU, Zheng  
; TITLE OF INVENTION: EPITOPE SEQUENCES  
; FILE REFERENCE: MANK.032VPC  
; CURRENT APPLICATION NUMBER: PCT/US03/27706  
; CURRENT FILING DATE: 2003-09-05  
; PRIOR APPLICATION NUMBER: US 60/409123  
; PRIOR FILING DATE: 2002-09-06  
; NUMBER OF SEQ ID NOS: 610  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 387  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homosapiens  
PCT-US03-27706-387

Query Match 38.0%; Score 19; DB 1; Length 10;  
Best Local Similarity 66.7%; Pred. No. 1.2e+03;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 6 EALAFI 11  
| | | |  
Db 1 ERLAYL 6

## RESULT 13

US-09-239-043D-19  
; Sequence 19, Application US/09239043D  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott

; APPLICANT: Vitiello, Maria A.  
; APPLICANT: Livingston, Brian D.  
; APPLICANT: Celis, Esteban  
; APPLICANT: Kubo, Ralph T.  
; APPLICANT: Grey, Howard M.  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Epimmune Inc.  
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Hepatitis B Virus  
; TITLE OF INVENTION: Using Peptide and Nucleic Acid Compositions  
; FILE REFERENCE: 2060.0060007  
; CURRENT APPLICATION NUMBER: US/09/239,043D  
; CURRENT FILING DATE: 1999-01-27  
; PRIOR APPLICATION NUMBER: US 09/189,702  
; PRIOR FILING DATE: 1998-11-10  
; PRIOR APPLICATION NUMBER: US 08/978,291  
; PRIOR FILING DATE: 1997-11-25  
; PRIOR APPLICATION NUMBER: US 08/820,360  
; PRIOR FILING DATE: 1997-03-12  
; PRIOR APPLICATION NUMBER: US 60/013,363  
; PRIOR FILING DATE: 1996-03-13  
; PRIOR APPLICATION NUMBER: US 08/461,603  
; PRIOR FILING DATE: 1995-06-05  
; PRIOR APPLICATION NUMBER: US 08/347,610  
; PRIOR FILING DATE: 1994-12-01  
; PRIOR APPLICATION NUMBER: US 08/344,824  
; PRIOR FILING DATE: 1994-11-23  
; PRIOR APPLICATION NUMBER: US 08/278,634  
; PRIOR FILING DATE: 1994-07-21  
; PRIOR APPLICATION NUMBER: US 08/205,713  
; PRIOR FILING DATE: 1994-03-04  
; PRIOR APPLICATION NUMBER: US 08/197,484  
; PRIOR FILING DATE: 1994-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 2579  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 19  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Orthohepadnaviridae hepatitis B virus  
US-09-239-043D-19

Query Match 38.0%; Score 19; DB 5; Length 10;  
Best Local Similarity 66.7%; Pred. No. 1.2e+03;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 6 EALAFI 11  
| | | |  
Db 1 ELLSFL 6

## RESULT 14

US-09-239-043D-696  
; Sequence 696, Application US/09239043D  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Vitiello, Maria A.  
; APPLICANT: Livingston, Brian D.  
; APPLICANT: Celis, Esteban  
; APPLICANT: Kubo, Ralph T.  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Epimmune Inc.  
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Hepatitis B Virus  
; TITLE OF INVENTION: Using Peptide and Nucleic Acid Compositions  
; FILE REFERENCE: 2060.0060007  
; CURRENT APPLICATION NUMBER: US/09/239,043D  
; CURRENT FILING DATE: 1999-01-27  
; PRIOR APPLICATION NUMBER: US 09/189,702  
; PRIOR FILING DATE: 1998-11-10  
; PRIOR APPLICATION NUMBER: US 08/978,291  
; PRIOR FILING DATE: 1997-11-25

; PRIOR APPLICATION NUMBER: US 08/820,360  
; PRIOR FILING DATE: 1997-03-12  
; PRIOR APPLICATION NUMBER: US 60/013,363  
; PRIOR FILING DATE: 1996-03-13  
; PRIOR APPLICATION NUMBER: US 08/461,603  
; PRIOR FILING DATE: 1995-06-05  
; PRIOR APPLICATION NUMBER: US 08/347,610  
; PRIOR FILING DATE: 1994-12-01  
; PRIOR APPLICATION NUMBER: US 08/344,824  
; PRIOR FILING DATE: 1994-11-23  
; PRIOR APPLICATION NUMBER: US 08/278,634  
; PRIOR FILING DATE: 1994-07-21  
; PRIOR APPLICATION NUMBER: US 08/205,713  
; PRIOR FILING DATE: 1994-03-04  
; PRIOR APPLICATION NUMBER: US 08/197,484  
; PRIOR FILING DATE: 1994-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 2579  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 696  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Orthohepadnaviridae hepatitis B virus  
US-09-239-043D-696

Query Match 38.0%; Score 19; DB 5; Length 10;  
Best Local Similarity 66.7%; Pred. No. 1.2e+03;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 6 EALAF1 11  
| : | |  
Db 1 ELLSFL 6

RESULT 15  
US-09-239-043D-1535  
; Sequence 1535, Application US/09239043D  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Vitiello, Maria A.  
; APPLICANT: Livingston, Brian D.  
; APPLICANT: Celis, Esteban  
; APPLICANT: Kubo, Ralph T.  
; APPLICANT: Grey, Howard M.  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Eplimmune Inc.  
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Hepatitis B Virus  
; TITLE OF INVENTION: Using Peptide and Nucleic Acid Compositions  
; FILE REFERENCE: 2060.0060007  
; CURRENT APPLICATION NUMBER: US/09/239,043D  
; CURRENT FILING DATE: 1999-01-27  
; PRIOR APPLICATION NUMBER: US 09/189,702  
; PRIOR FILING DATE: 1998-11-10  
; PRIOR APPLICATION NUMBER: US 08/978,291  
; PRIOR FILING DATE: 1997-11-25  
; PRIOR APPLICATION NUMBER: US 08/820,360  
; PRIOR FILING DATE: 1997-03-12  
; PRIOR APPLICATION NUMBER: US 60/013,363  
; PRIOR FILING DATE: 1996-03-13  
; PRIOR APPLICATION NUMBER: US 08/461,603  
; PRIOR FILING DATE: 1995-06-05  
; PRIOR APPLICATION NUMBER: US 08/347,610  
; PRIOR FILING DATE: 1994-12-01  
; PRIOR APPLICATION NUMBER: US 08/344,824  
; PRIOR FILING DATE: 1994-11-23  
; PRIOR APPLICATION NUMBER: US 08/278,634  
; PRIOR FILING DATE: 1994-07-21  
; PRIOR APPLICATION NUMBER: US 08/205,713  
; PRIOR FILING DATE: 1994-03-04  
; PRIOR APPLICATION NUMBER: US 08/197,484  
; PRIOR FILING DATE: 1994-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 2579  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1535  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Orthohepadnaviridae hepatitis B virus  
US-09-239-043D-1535  
  
Query Match 38.0%; Score 19; DB 5; Length 10;  
Best Local Similarity 66.7%; Pred. No. 1.2e+03;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 6 EALAF1 11  
| : | |  
Db 1 ELLSFL 6  
  
Search completed: October 7, 2003, 14:24:30  
Job time : 25 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:16:45 ; Search time 57.5 Seconds  
(without alignments)  
27.605 Million cell updates/sec

Title: US-09-807-512-12  
Perfect score: 45  
Sequence: 1 LMAQEAALFL 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 251420

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :				
A_Geneseq_19Jun03:*				
1:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*			
2:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*			
3:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*			
4:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*			
5:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*			
6:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*			
7:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*			
8:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*			
9:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*			
10:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*			
11:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*			
12:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*			
13:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*			
14:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*			
15:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*			
16:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*			
17:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*			
18:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*			
19:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*			
20:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*			
21:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*			
22:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*			
23:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*			
24:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*			

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	100.0	10	21	AA1980
2	41	91.1	10	23	ABG75072
3	27	60.0	9	23	AB879090
4	24	53.3	9	13	AA28520
5	24	53.3	9	15	AA46863
6	24	53.3	9	15	AA55670
7	24	53.3	9	17	AA86655
8	24	53.3	10	16	AA867176
9	24	53.3	10	17	AAW03313

CTL epitope-2 of h  
Human CAMEL class  
Human tumour antig  
Novel peptide GGF-  
Cell proliferation  
30-36 kd glial gro  
GGF I tryptic pept  
Bovine glial cell  
Tyrosinase peptide

10	24	53.3	10	17	AA86642	GGF I tryptic pept
11	24	53.3	10	23	AAU71255	Human MHC class I
12	22	48.9	8	23	ABB79903	Human BM40 basemen
13	22	48.9	9	20	AAW84253	Antigenic peptide
14	22	48.9	9	20	AAW84219	Antigenic peptide
15	22	48.9	9	20	AAW84220	Antigenic peptide
16	22	48.9	9	21	AAU70857	CAMEL10 immunogeni
17	22	48.9	9	21	AAU70857	HLA-A24-binding HI
18	22	48.9	9	22	AAG99440	PRAME derived HLA-
19	22	48.9	9	22	AAG99483	PRAME derived HLA-
20	22	48.9	9	23	AAU71531	Human MHC class I
21	22	48.9	9	23	AAU71606	Human MHC class I
22	22	48.9	9	23	AAU71739	Human MHC molecule
23	22	48.9	10	22	AAG99456	PRAME derived HLA-
24	21	46.7	9	17	AAW01006	19K apoptosis-regu
25	21	46.7	9	20	AAW94331	19K binding intera
26	21	46.7	9	20	AAW84221	Antigenic peptide
27	21	46.7	9	23	ABB96933	Human tumour anti
28	21	46.7	9	23	AAU71132	Human MHC class I
29	21	46.7	9	23	AAU71502	Human MHC class I
30	21	46.7	9	23	AAU71726	Human MHC molecule
31	21	46.7	9	24	ABJ19933	MHC binding peptid
32	21	46.7	9	24	ABP74862	Proteome analysis
33	21	46.7	10	19	AAW41760	Colony stimulating
34	21	46.7	10	22	AAG96587	Human complementar
35	21	46.7	10	22	AAG83629	Arabidopsis thalia
36	21	46.7	10	22	AA85878	O. vulgaris tachyk
37	21	46.7	10	22	ABG98784	F protein decapept
38	21	46.7	10	23	ABB96935	Human tumour anti
39	21	46.7	10	23	ABB96936	Human tumour anti
40	21	46.7	10	23	AAU71156	Human MHC class I
41	21	46.7	10	23	AAU71171	Human MHC class I
42	21	46.7	10	23	AAU71490	Human MHC molecule
43	20	44.4	8	8	AAU70001	cdNA obtained foll
44	20	44.4	8	23	ABB80064	Factor VII fragmen
45	20	44.4	9	17	AAW03312	Tyrosinase peptide

ALIGNMENTS

RESULT 1

AAU70856  
ID AAU70856 standard; peptide; 10 AA.

XX AAU70856;

AC AAU70856;

XX 31-JUL-2000 (first entry)

XX CTL epitope-2 of human CAMEL protein.

XX CAMEL; CTL-recognised Antigen on MELanoma; cytotoxic T lymphocyte; CTL;

XX tumour-associated antigen; LAGE-I; NY-ESO-1; anticancer; melanoma;

XX human; cancer; immunotherapy; immunogenic peptide; immune response.

XX Homo sapiens.

XX WO200023584-A1.

XX 27-APR-2000.

XX 15-OCT-1999; 99WO-EP07832.

XX 16-OCT-1998; 98EP-0119583.

XX (BOEH ) BOEHRINGER INGELHEIM INT GMBH.

XX (UYHO-) UNIV HOSPITAL LEIDEN.

XX Schrier PI, Aarnoudse CA, Heider K, Klade C;

XX WPI; 2000-339685/29.

XX Tumor-associated antigen useful for cancer immunotherapy is encoded by

PT the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA -

XX Claim 5; Page 34; 73pp; English.

XX The present sequence is an immunogenic peptide of human tumour-associated

CC antigen CAMEL (Cytotoxic T lymphocytes (CTL)-recognised Antigen on

CC MELANOMA). This peptide is a CTL epitope, that has the ability to elicit

CC a CTL response. It corresponds to residues 2-11 of the CAMEL protein.

CC CAMEL protein is encoded by the LAGE-1 gene, a tumour-specific antigen.

CC It is different from the LAGE-1 protein, since it is translated from a

CC different open reading frame (ORF-1). It shows strong homology with

CC NY-ESO-1, a melanoma specific tumour antigen. The tumour-associated

CC antigen displayed on melanoma cells is recognised by cytotoxic T

CC lymphocytes. CAMEL is expressed in tumour cell lines, tumour tissues

CC (e.g. breast and lung) and in restricted number of healthy tissues. This

CC sequence has anticancer activity. CAMEL tumour antigen and immunogenic

CC peptides derived from it are useful for cancer immunotherapy. They have

CC the potential to induce an immune response, by eliciting a CTL response.

CC The DNA molecule is used to construct recombinant or fusion proteins.

XX Sequence 10 AA;

SQ Query Match 100.0%; Score 45; DB 21; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.022;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LMAQEALAF 10

Db 1 LMAQEALAF 10

RESULT 2

ABG79072

ID ABG79072 standard; Peptide; 10 AA.

AC ABG79072;

XX 15-NOV-2002 (first entry)

DT Human CAMEL class I HLA widely expressed antigen peptide #1.

DE

XX Cell penetrating peptide; cancer; tumour; melanoma; thymoma; antigen;

KW lymphoma; sarcoma; lung cancer; non-Hodgkin's lymphoma; leukaemia;

KW Hodgkin's lymphoma; uterine cancer; cervical cancer; bladder cancer;

KW kidney cancer; adenocarcinoma; breast cancer; prostate cancer;

KW ovarian cancer; pancreatic cancer; epitope; vaccine; dendritic cell;

KW tumour infiltrating lymphocyte; TIL; human leukocyte antigen; HLA;

KW cytostatic; human.

XX Homo sapiens.

OS WO200264057-A2.

XX PN 22-AUG-2002.

XX 15-FEB-2002; 2002WO-US05212.

XX PF 15-FEB-2001; 2001US-268687P.

XX PR (BAYU ) BAYLOR COLLEGE MEDICINE.

XX PA Wang R;

XX PI WPI; 2002-627577/67.

XX DR Novel composition for treating a disease in an animal, comprises an

XX PT immune effector cell and cell penetrating peptide associated with an

PT antigen or antibody -

XX Disclosure; Page 17; 61pp; English.

PS The invention relates to a composition (I) comprising an immune effector

XX cell and a cell penetrating peptide (CPP) associated with an antigen or

CC

CC antibody. Also included are (1) a vaccine comprising (I), CPP

CC associated with an antigen, and a pharmaceutically acceptable carrier

CC and (2) preparing a composition for a disease, by providing (I)

CC and CPP associated with an antigen for disease, and introducing the

CC antigen-associated CPP to (I), where antigen enters into the cell.

CC The antigens are, for example, tumour antigen derived epitopes

CC recognised by tumour infiltrating lymphocytes (TIL) of HLA (human

CC leukocyte antigen) class I or II. The composition is useful for enhancing

CC immunity in an animal to a disease, by administering a mature dendritic

CC cell comprising CPP associated with an antigen to disease, to the animal,

CC such that following the administration, animal is protected from disease,

CC where the animal comprises both CD4+ and CD8+ T cells. It is also useful

CC for treating a disease (e.g. cancer, tumour, melanoma, thymoma,

CC lymphoma, sarcoma, lung cancer, non-Hodgkin's lymphoma, leukaemia,

CC Hodgkin's lymphoma, uterine cancer, cervical cancer, bladder cancer,

CC kidney cancer, adenocarcinoma, breast cancer, prostate cancer,

CC ovarian cancer and pancreatic cancer). The animal is further subjected to

CC a cancer treatment including surgery, radiation, chemotherapy or gene

CC therapy. The administration of (I), preferably dendritic cell is prior

CC to, subsequent to or concurrent with, the cancer treatment. The present

CC sequence is a tumour antigen derived epitope for inclusion in the

CC composition of the invention.

XX Query Match 91.1%; Score 41; DB 23; Length 10;

SQ Best Local Similarity 100.0%; Pred. No. 0.13;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LMAQEALAF 9

Db 2 LMAQEALAF 10

RESULT 3

ABB97090

ID ABB97090 standard; Peptide; 9 AA.

XX AC ABB97090;

XX 21-JUN-2002 (first entry)

DT Human tumour antigen related peptide 141.

XX Human; tumour; antigen; HLA-A2; cytotoxic T cell; cytostatic; cancer;

KW vaccine.

KW Homo sapiens.

XX OS WO200210369-A1.

XX PN 07-FEB-2002.

XX PD 30-JUL-2001; 2001WO-JP06526.

XX PF 31-JUL-2000; 2000JP-0231814.

XX PR (ITOH/) ITOH K.

XX PA Itoh K;

XX PI WPI; 2002-291857/33.

XX DR Tumor antigens inducing and/or activating HLA-A2-restricted

XX PT tumor-specific cytotoxic T cells, useful in diagnosis of and screening

PT drugs e.g. cancer vaccines for specific treatment of pancreatic cancer

PT

XX Example 4; Fig 10; 127pp; Japanese.

PS The invention relates to a peptide comprising an amino acid sequence

XX selected from 44 fully defined amino acid sequences (ABB96906-ABB96949)

CC and a polypeptide comprising an amino acid sequence selected from the 9

CC

CC fully defined amino acid sequences (ABB97143-ABB97151). The above  
 CC comprise a tumour antigen inducing or activating HLA-A2-restricted  
 CC tumour-specific cytotoxic T cells, which recognise HLA-A2 and a tumour  
 CC antigen peptide and is thus activated. The peptides and polypeptides  
 CC have cytostatic activity. The tumour antigen is useful in diagnosis of  
 CC and screening drugs for specific treatment of pancreatic cancer, colon  
 CC cancer and stomach cancer including in the form of vaccines. The present  
 CC sequence is that of a tumour antigen peptide, useful to the invention.  
 SQ Sequence 9 AA;

Query Match 60.0%; Score 27; DB 23; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 9.3e+05;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMAQEALA 8  
 I:|||||  
 DB 1 LLAQEAAA 8

RESULT 4  
 AAR28520  
 ID AAR28520 standard; Protein; 9 AA.  
 XX  
 AC AAR28520;

XX 25-MAR-2003 (updated)  
 DT 18-MAR-1993 (first entry)

DE Novel peptide GGF-I 14.

XX Glial growth factor-II; GGF-II; backtranslation; bovine cDNA library;  
 KW polymerase chain reaction; PCR; amplify; pituitary RNA; exons;  
 KW alternative splice pattern.

XX Synthetic.

XX WO9218627-A1.

XX 29-OCT-1992.

XX 03-APR-1992; 92WO-GB00595.

XX 10-APR-1991; 91GB-0007566.

XX (CAMP-) CAMBRIDGE NEUROSCIENCE INC.  
 PA (LUDW-) LUDWIG INST CANCER RES.

PI Goodearl ADJ, Marchionni MA, Minghetti L, Stroobant P;  
 PI Waterfield MD;

XX WPI; 1992-382111/46.

XX Glial mitogenic polypeptide factors - used in diagnosis,  
 PT prophylaxis or treatment of patho-physiological conditions of the  
 PT nervous system

PS Disclosure; Fig 10; 98pp; English.

XX The sequences given in AAR28514-25 are peptides derived from bovine  
 CC glial growth factor-I (GGF-I). Backtranslation of these sequences may  
 CC provide probes and primers which can be used for the isolation of the  
 CC GGF-I cDNA. A bovine cDNA library may be probed with the  
 CC backtranslated probes. Isolated open reading frames may then be  
 CC amplified using polymerase chain reaction on pituitary RNA.  
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 9 AA;

Query Match 53.3%; Score 24; DB 13; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 9.3e+05;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 AQEALAF 10  
 I:|||||  
 DB 1 AKEALAA 8

RESULT 5  
 AAR46863  
 ID AAR46863 standard; Protein; 9 AA.  
 XX  
 AC AAR46863;

XX 25-MAR-2003 (updated)  
 DT 10-MAR-2003 (updated)  
 DT 28-JUL-1994 (first entry)

XX Cell proliferation inhibitor.

XX Glial growth factor; GGF; heregulin; mitogenesis;  
 KW Schwann cell; tumour; central nervous system;  
 KW antiproliferative; inhibitor.

XX Unidentified.

XX WO9403644-A1.

XX 17-FEB-1994.

XX 10-AUG-1993; 93WO-US07491.

XX 10-AUG-1992; 92US-0927337.

XX 25-SEP-1992; 92US-0951747.

XX 01-DEC-1992; 92US-0984085.

XX 29-JAN-1993; 93US-0011396.

XX (CAMP-) CAMBRIDGE NEUROSCIENCE INC.

XX Gwynne DI, Marchionni M, McBurney RN;

XX WPI; 1994-065731/08.

XX Glial growth factor DNA encoding numerous polypeptide factors  
 PT used for inhibiting cell proliferation - for treating carcinoma  
 PT and nervous disorders

XX Disclosure; Fig 10; 178pp; English.

XX Factors useful in the methods of the invention are  
 CC basic polypeptide factors having antiproliferative  
 CC activity when contacted with cells, including cells of the  
 CC nervous system and specifically Schwann cells, and contg.  
 CC within their amino acid sequences at least one of the polypeptides  
 CC given in AAR46857-68.

XX (Updated on 10-MAR-2003 to add missing OS field.)

XX (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 9 AA;

Query Match 53.3%; Score 24; DB 15; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 9.3e+05;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 AQEALAF 10  
 I:|||||  
 DB 1 AKEALAA 8

RESULT 6  
 AAR55670  
 ID AAR55670 standard; Protein; 9 AA.  
 XX  
 AC AAR55670;

XX 25-MAR-2003 (updated)





XX PF 06-MAY-1994; 94WO-US05083.  
 XX PR 06-MAY-1993; 93US-0059022.  
 XX PR 08-MAR-1994; 94US-0209204.  
 XX PA (CAMB-) CAMBRIDGE NEUROSCIENCE.  
 XX PI Gwynne DI, Marchionni M, Sklar R;  
 XX PI WPI; 1995-006353/01.  
 XX DR  
 XX PT Treating mammalian muscle diseases and disorders - by admin. of  
 XX PT GGF2 and other specified polypeptide(s) which bind the p185erbB2  
 XX PT receptor.  
 XX PS Example 9; Page 91; 241pp; English.  
 XX XX  
 CC AAR67153-R67163 and AAR67174-R67183 are bovine glial cell growth  
 CC factor-1 (GGF-1) peptide fragments. A 30-36kD polypeptide factor  
 CC known to have glial cell mitogenic activity, which includes one  
 CC of the above peptide fragments in its amino acid sequence is  
 CC claimed. The polypeptide can be used to treat a variety of  
 CC mammalian skeletal, cardiac and smooth muscle diseases,  
 CC including acetylcholine receptor deficiency.  
 CC CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX AC  
 XX SQ Sequence 10 AA;  
 Query Match 53.3%; Score 24; DB 16; Length 10;  
 Best Local Similarity 75.0%; Pred. No. 2.9e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 AQEALAEFL 10  
 DB 2 AKELAAAL 9  
 RESULT 9  
 AAW03313  
 ID AAW03313 standard; peptide; 10 AA.  
 XX AC AAW03313;  
 XX XX  
 XX DT 22-OCT-1996 (first entry)  
 XX DE Tyrosinase peptide T10199 (aa199-208).  
 XX KW Melanoma; tyrosinase; immunogen; vaccine; cancer; immunotherapy.  
 XX OS Synthetic.  
 XX XX  
 XX PN WO9621734-A2.  
 XX PD 18-JUL-1996.  
 XX PF 11-JAN-1996; 96WO-US00473.  
 XX PR 10-JAN-1995; 95US-0370909.  
 XX XX  
 XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.  
 XX PI Robbins PF, Rosenberg SA;  
 XX PI WPI; 1996-342287/34.  
 XX XX  
 CC Nucleic acid sequence encoding p15 melanoma antigen - and  
 CC immunogenic peptide(s) derived from it, useful for diagnosis,  
 CC prevention or treatment of melanoma  
 XX XX  
 XX PS Example 2; Page 64; 98pp; English.  
 XX CC A series of synthetic peptides (AAW03307-15 plus AAW03304-05) are

CC based on putative peptide binding motifs to HLA-A24 in human  
 CC tyrosinase melanoma antigen (AAW03306). Peptide AAT10199 (AAW03313)  
 CC corresponds to amino acids 199-208 of tyrosinase. The peptides  
 CC were tested for their ability to sensitize HLA-A24+ 888  
 CC Epstein-Barr virus-transformed B-cells to tumour infiltrating  
 CC lymphocyte (TIL) lysis and their ability to render melanoma 888  
 CC cells to stimulate GM-CSF release from TIL. Positive results  
 CC were obtd. with peptides T9206 (AAW03304) and AAT10206 (AAW03305).  
 CC These peptides can be used in melanoma vaccines.  
 XX XX  
 XX SQ Sequence 10 AA;  
 Query Match 53.3%; Score 24; DB 17; Length 10;  
 Best Local Similarity 75.0%; Pred. No. 2.9e+02;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 3 AQEALAEFL 10  
 DB 3 AHEAPAEFL 10  
 RESULT 10  
 AAR86642  
 ID AAR86642 standard; peptide; 10 AA.  
 XX AC AAR86642;  
 XX XX  
 XX DT 03-JUL-1996 (first entry)  
 XX DE GGF I tryptic peptide GGF-I 14 #1.  
 XX KW Glial growth factor; GGF; human; hGGF2; Schwann cell; mitogenesis; GGF-I;  
 XX KW glial cell; therapy; peripheral nerve damage; demyelination; bovine;  
 XX KW glial tumour; fibroblast proliferation; wound repair; multiple sclerosis;  
 XX KW neurodegenerative disorder; neural regeneration; acetylcholine receptor.  
 XX OS Synthetic.  
 XX XX  
 XX FH Key Location/Qualifiers  
 XX FT Misc-difference 1  
 XX FT /label= Lys, Arg  
 XX PN WO9532724-A1.  
 XX PD 07-DEC-1995.  
 XX PF 25-MAY-1995; 95WO-US06846.  
 XX PR 26-MAY-1994; 94US-0249322.  
 XX XX  
 XX PA (CAMB-) CAMBRIDGE NEUROSCIENCE.  
 XX PA (LUDW-) LUDWIG INST CANCER RES.  
 XX PI Chen MS, Goodearl AD, Hiles I, Marchionni M, Minghetti L;  
 XX PI Stroobant P, Waterfield M;  
 XX XX  
 XX WPI; 1996-030329/03.  
 XX DR N-PSDB; AAT06646, AAT06647, AAT06667, AAT06688, AAT06690.  
 XX XX  
 XX PT Glial growth factors with Schwann cell mitogenic activity - for  
 XX PT prophylaxis or treatment of nerve disorders, e.g. Multiple Sclerosis  
 XX PS Example 2; Fig 9; 199pp; English.  
 XX XX  
 CC AAR86629-R86657 represent fragments of bovine glial growth factor I  
 CC (GGF-I). These fragments were obtained by lysyl endopeptidase and  
 CC protease V8 digests. These sequences have Schwann cell mitogenic  
 CC activity in the presence of foetal calf plasma. These sequences (and  
 CC human GGF2, see AAR86628) are used to stimulate mitogenesis of glial  
 CC cells, for prophylaxis or treatment of a pathological condition of  
 CC the nervous system in a mammal. Also, for identification of a receptor,  
 CC for treatment of conditions of peripheral nerve damage (e.g.  
 CC demyelination/damage/loss of Schwann cells), treatment of

CC neurodegenerative disorders in mammals, for inducing neural  
CC regeneration, fibroblast proliferation or wound repair. Glial tumours  
CC can be treated by inhibiting this sequence binding to its receptor.  
CC These peptides are also able to induce synthesis of acetylcholine  
CC receptor, and is useful in prophylaxis or treatment of multiple  
CC sclerosis in a patient.  
XX  
SQ Sequence 10 AA;  
  
Query Match 53.3%; Score 24; DB 17; Length 10;  
Best Local Similarity 75.0%; Pred. No. 2.9e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 3 AQEALAF 10  
DB 2 AQEALAF 9  
  
RESULT 11  
AAU71255  
ID AAU71255 standard; Peptide; 10 AA.  
XX AAU71255;  
XX  
XX 26-FEB-2002 (first entry)  
XX Human MHC class I molecule HLA-A2 binding 103P3E8 peptide #62.  
XX 103P3E8; prostate; bladder; kidney; lung; breast; rectum; stomach;  
KW tumour; cancer; cytostatic; gene therapy; antibody therapy; ribozyme;  
KW single chain monoclonal antibody; serum; blood; urine; tissue; human;  
KW chromosome 9q13-q21.  
XX  
XX Homo sapiens.  
XX WO200179557-A2.  
XX  
XX 25-OCT-2001.  
XX 12-APR-2001; 2001WO-US12181.  
XX 12-APR-2000; 2000US-196647P.  
XX  
XX (UROC-) UROGENESYS INC.  
XX Faris M, Challita-Eid PM, Raitano AB, Mitchell SC, Afar DEH;  
PI Jakobovits A;  
XX  
XX WPI; 2002-061976/08.  
XX  
XX Monitoring 103P3E8 gene products in sample from patient (suspected of)  
XX having cancer, useful for diagnosing, managing or treating cancers,  
XX e.g. prostate cancer, comprises determining presence of aberrant  
XX 103P3E8 gene products -  
XX  
XX Disclosure; Page 85; 128pp; English.  
XX  
XX Sequences AAU71093-AAU71796 represent the 103P3E8-related protein and  
XX peptide fragments of the protein. 103P3E8 exhibits tissue specific  
XX expression in normal adult tissue, but it is also aberrantly expressed in  
XX many cancers including tumours of the prostate, bladder, kidney, colon,  
XX lung, breast, rectum and stomach. The 103P3E8 polynucleotide, its related  
XX protein and peptide fragments and specific PCR primers are therefore  
XX useful for diagnosing and treating cancer. A vector comprising a  
XX polynucleotide which encodes a single chain monoclonal antibody, that  
XX immunospecifically binds to an 103P3E8-related protein, and a ribozyme  
XX capable of cleaving a polynucleotide having the 103P3E8 coding sequence,  
XX are both useful in the preparation of a composition for treating a  
XX patient with a cancer that expresses 103P3E8. The sequences can be used  
XX in diagnostic methods to monitor the level of 103P3E8 gene products in  
XX serum, blood, urine and tissue and to thereby detect the presence of  
XX cancerous cells.

SQ Sequence 10 AA;  
  
Query Match 53.3%; Score 24; DB 23; Length 10;  
Best Local Similarity 60.0%; Pred. No. 2.9e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 LMAQALAF 10  
DB 1 LEAQTNI 10  
  
RESULT 12  
ABB79903  
ID ABB79903 standard; Peptide; 8 AA.  
XX ABB79903;  
XX  
XX 05-DEC-2002 (first entry)  
XX Human BM40 basement signal peptide.  
XX Endostatin; human; ophthalmological; ocular neovascularisation;  
KW choroidal neovascularisation; gene therapy; BM40; signal peptide.  
XX Homo sapiens.  
XX WO200267971-A2.  
XX  
XX 06-SEP-2002.  
XX 21-FEB-2002; 2002WO-US05336.  
XX 22-FEB-2001; 2001US-270787P.  
XX 04-APR-2001; 2001US-281296P.  
XX  
XX (NOVS ) NOVARTIS AG.  
XX  
XX Brazzell RK, Campochiaro PA, Dixon KH;  
XX WPI; 2002-698636/75.  
XX  
XX Treating or preventing choroidal neovascularization comprises  
XX increasing the amount of endostatin in ocular tissues of afflicted  
XX individuals to a choroidal neovascularization inhibiting level -  
XX  
XX Example 6; Page 29; 44pp; English.  
XX  
XX The present sequence comprises a portion of the human BM40 basement  
XX protein signal peptide. In an example from the invention, a  
XX recombinant adenoviral vector encoding human endostatin plus the  
XX BM40 basement protein leader was generated. Vector-transduced g8  
XX cells produced 3 major secreted proteins, 50% containing the amino  
XX acid sequence (see ABB79901) of human endostatin with the present  
XX additional amino acid residues, 25% containing residues LA, and 25%  
XX containing no residues from BM40 signal peptide. A claimed method  
XX for the treatment of ocular neovascularisation, especially choroidal  
XX neovascularisation, involves increasing the level of endostatin in  
XX ocular tissue. The increase is effected by administering a viral  
XX vector, especially an adenovirus, adeno-associated virus, a  
XX retrovirus or lentivirus vector, comprising an endostatin-encoding  
XX nucleic acid. Cells secreting endostatin may be encapsulated and  
XX implanted within an individual.  
  
SQ Sequence 8 AA;  
  
Query Match 48.9%; Score 22; DB 23; Length 8;  
Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 4 QEALA 8  
DB 4 QEALA 8

XX	GP4 protein; PRRSV; structural protein; neutralising antibody; antigen;
KW	vaccine; pig; PRRSV infection; isolate detection.
XX	Porcine reproductive and respiratory syndrome virus.
OS	
XX	WO9850426-A1.
PN	
XX	12-NOV-1998.
PD	
XX	
XX	05-MAY-1998; 98WO-NL00251.
PF	
XX	
PR	06-MAY-1997; 97EP-0201343.
XX	(DIER-) STICHTING INST DIERHOUDERIJ EN DIERGEZON.
PA	Langeveld J, Meulenbergh J, Van Nieuwstadt AP;
XX	WPI; 1999-070090/06.
PI	
XX	New peptides containing epitopes of porcine reproductive and
PT	respiratory syndrome virus - useful in vaccines and for diagnostic
XX	tests, particularly for differentiating between infected and
PT	vaccinated animals
XX	
PS	Claim 6; Page 29; 46pp; English.
XX	The present sequence represents a peptide derived from the GP4 protein
CC	(encoded by open reading frame 4 of the virion) of Porcine reproductive
CC	and respiratory syndrome virus (PRRSV). GP4 is a structural protein.
CC	The present peptide is able to elicit neutralising antibodies. Such
CC	antigenic peptides are used in vaccines to protect pigs against PRRSV
CC	infection, thus reducing the occurrence of PRRSV in pig herds and aiding
CC	its eradication. The peptides may also be used diagnostically for
CC	detection of PRRSV-specific antibodies, i.e. to diagnose infection, to
CC	differentiate between vaccinated and naturally infected animals and to
CC	distinguish between different PRRSV isolates. Synthetic antibodies
CC	are used to detect antigens containing the epitopes present in the
CC	peptides.
XX	
SQ	Sequence 9 AA;
	Query Match 48.9%; Score 22; DB 20; Length 9;
	Best Local Similarity 57.1%; Pred. No. 9.3e+05;
	Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY	3 AQEALAF 9    :::
Dd	3 AQEKISF 9
	RESULT 15
	AAW84220
ID	AAW84220 standard; peptide; 9 AA.
XX	
AC	AAW84220;
XX	
DT	25-MAR-1999 (first entry)
XX	
DE	Antigenic peptide derived from the GP4 protein of PRRSV.
XX	
KW	GP4 protein; PRRSV; structural protein; neutralising antibody; antigen;
KW	vaccine; pig; PRRSV infection; isolate detection.
XX	Porcine reproductive and respiratory syndrome virus.
OS	
XX	WO9850426-A1.
PN	
XX	12-NOV-1998.
PD	
XX	
PF	05-MAY-1998; 98WO-NL00251.
XX	
PR	06-MAY-1997; 97EP-0201343.
XX	(DIER-) STICHTING INST DIERHOUDERIJ EN DIERGEZON.
PA	Langeveld J, Meulenbergh J, Van Nieuwstadt AP;
XX	WPI; 1999-070090/06.
PI	
XX	New peptides containing epitopes of porcine reproductive and
PT	respiratory syndrome virus - useful in vaccines and for diagnostic
XX	tests, particularly for differentiating between infected and
PT	vaccinated animals
XX	
PS	Claim 5; Page -; 46pp; English.
XX	The present sequence represents a peptide derived from amino acids
CC	59 to 67 of the GP4 protein (encoded by open reading frame 4 of the
CC	virion) of Porcine reproductive and respiratory syndrome virus
CC	(PRRSV) isolate I-1102. GP4 is a structural protein. The present
CC	peptide is able to elicit neutralising antibodies. Such antigenic
CC	peptides are used in vaccines to protect pigs against PRRSV infection,
CC	thus reducing the occurrence of PRRSV in pig herds and aiding in its
CC	eradication. The peptides may also be used diagnostically for
CC	detection of PRRSV-specific antibodies, i.e. to diagnose infection, to
CC	differentiate between vaccinated and naturally infected animals and to
CC	distinguish between different PRRSV isolates. Synthetic antibodies
CC	are used to detect antigens containing the epitopes present in the
CC	peptides.
CC	note: this sequence does not appear in the specification; it was created
CC	using information provided.
XX	
SQ	Sequence 9 AA;
	Query Match 48.9%; Score 22; DB 20; Length 9;
	Best Local Similarity 57.1%; Pred. No. 9.3e+05;
	Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY	3 AQEALAF 9    :::
Dd	3 AQEKISF 9
	RESULT 14
	AAW84219
ID	AAW84219 standard; peptide; 9 AA.
XX	
AC	AAW84219;
XX	
DT	25-MAR-1999 (first entry)
XX	
DE	Antigenic peptide derived from the GP4 protein of PRRSV.

XX (DIER-) STICHTING INST DIERHOUDERIJ EN DIERGEZON.  
XX Langeveld J, Meulenberg J, Van Nieuwstadt AP;  
XX WPI; 1999-070090/06.  
XX New peptides containing epitopes of porcine reproductive and  
PT respiratory syndrome virus - useful in vaccines and for diagnostic  
PT tests, particularly for differentiating between infected and  
PT vaccinated animals  
XX  
XX Claim 6; Page 29; 46pp; English.  
XX  
CC The present sequence represents a peptide derived from the GP4 protein  
CC (encoded by open reading frame 4 of the virion) of Porcine reproductive  
CC and respiratory syndrome virus (PRRSV). GP4 is a structural protein.  
CC The present peptide is able to elicit neutralising antibodies. Such  
CC antigenic peptides are used in vaccines to protect pigs against PRRSV  
CC infection, thus reducing the occurrence of PRRSV in pig herds and aiding  
CC in its eradication. The peptides may also be used diagnostically for  
CC detection of PRRSV-specific antibodies, i.e. to diagnose infection, to  
CC differentiate between vaccinated and naturally infected animals and to  
CC distinguish between different PRRSV isolates. Synthetic antibodies  
CC are used to detect antigens containing the epitopes present in the  
CC peptides.  
XX  
SQ Sequence 9 AA;  
  
Query Match 48.9%; Score 22; DB 20; Length 9;  
Best Local Similarity 57.1%; Pred. No. 9.3e+05;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
QY 3 AQEALAF 9  
   | | | : : |  
Db 3 AQEEISF 9  
  
Search completed: October 7, 2003, 14:27:44  
Job time : 58.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:24:35 ; Search time 24 seconds  
(without alignments)  
40.070 Million cell updates/sec

Title: US-09-807-512-12  
Perfect score: 45  
Sequence: 1 LMAQELAF 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues  
Total number of hits satisfying chosen parameters: 1100

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_76:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	40.0	9	2 A33527	fructose-2,6-bisph
2	15	33.3	9	2 A61230	calsequestrin, car
3	15	33.3	10	2 S70721	heat shock protein
4	14	31.1	6	2 A37765	hypothetical prote
5	14	31.1	8	2 S6296	Na+-transporting A
6	14	31.1	9	2 D28854	fibrinopeptide B -
7	14	31.1	9	2 E28854	fibrinopeptide B -
8	14	31.1	9	2 F28854	fibrinopeptide B -
9	14	31.1	9	2 JN0027	[Phe-6]-mosact - s
10	14	31.1	10	2 S24190	tryptase (EC 3.4.2
11	14	31.1	10	2 A60722	cryptic fibrillar p
12	14	31.1	10	2 PH0113	alpha-amylin (EC
13	13	28.9	7	2 P0246	Ig heavy chain CRD
14	13	28.9	8	2 S6646	cardioacceleratory
15	13	28.9	8	2 S65381	cytochrome-c oxida
16	13	28.9	9	2 B41978	calliFMRFamide 2 -
17	13	28.9	9	2 C41978	calliFMRFamide 3 -
18	13	28.9	10	2 S28055	cytochrome b559 co
19	12	26.7	4	2 S18401	throglobulin - do
20	12	26.7	6	2 T11779	phosphoglycerate t
21	12	26.7	6	4 S15596	orf 3 rara 5'-regi
22	12	26.7	7	2 S55548	mcrB protein - Esc
23	12	26.7	7	2 PQ0777	NADH2 dehydrogenas
24	12	26.7	8	2 A37521	R-phycocerythrin ga
25	12	26.7	8	2 I48935	apolipoprotein A-I
26	12	26.7	8	2 B47594	aspartate kinase (
27	12	26.7	8	2 S33008	citrate synthase -
28	12	26.7	9	2 C24180	fibrinogen beta ch
29	12	26.7	9	2 S36898	ribosomal protein

30	12	26.7	9	2 A41978	calliFMRFamide 1 -
31	12	26.7	9	2 D41978	calliFMRFamide 4 -
32	12	26.7	9	2 E41978	calliFMRFamide 5 -
33	12	26.7	9	2 F41978	calliFMRFamide 6 -
34	12	26.7	9	2 G41978	calliFMRFamide 7 -
35	12	26.7	10	2 P0136	pepsin (EC 3.4.23.
36	12	26.7	10	2 S70251	nitrogenase (EC 1.
37	12	26.7	10	2 I40032	trpE protein - Bac
38	12	26.7	10	2 I44644	neurotoxin-associa
39	12	26.7	10	2 D33098	214K exoantigen (v
40	12	26.7	10	2 A58365	neuropeptide FFRfa
41	12	26.7	10	2 T14212	cytochrome-c oxida
42	12	26.7	10	2 PQ0783	NADH2 dehydrogenas
43	12	26.7	10	4 S14943	UGA3 leader peptid
44	11	24.4	5	2 B41225	copper resistance
45	11	24.4	5	2 I40469	dnazX-like protein

ALIGNMENTS

RESULT 1  
A33527 fructose-2,6-bisphosphate 2-phosphatase (EC 3.1.3.46) - rat (fragment)  
N:Alternate names: fructose-2,6-bisphosphatase  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 27-Feb-1990 #sequence\_revision 27-Feb-1990 #text\_change 23-Jun-1993  
C:Accession: A33527  
R:Kitamura, K.; Uyeda, K.; Hartman, F.C.; Kangawa, K.; Matsuo, H.  
J. Biol. Chem. 264, 6344-6348, 1989  
A:Title: Catalytic site of rat liver and bovine heart fructose-6-phosphate,2-kinase:f  
A:Reference number: A33527; MUID:89197937; PMID:2539378  
A:Accession: A33527  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-9 <KIT>  
C:Keywords: phosphoric monoester hydrolase

Query Match 40.0%; Score 18; DB 2; Length 9;  
Best Local Similarity 71.4%; Pred. No. 2.8e+05;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 QEALAF 10  
| | | | |  
Db 2 QXALAL 8

RESULT 2  
A61230 calsequestrin, cardiac and slow skeletal muscle - northern leopard frog (fragment)  
N:Alternate names: 58K dihydropyridine-binding protein; aspartactin; calmitine; lamin  
C:Species: Rana pipiens (northern leopard frog)  
C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 16-Jul-1999  
C:Accession: A61230  
R:McLeod, A.G.; Shen, A.C.Y.; Campbell, K.P.; Michalak, M.; Jorgensen, A.O.  
Circ. Res. 69, 344-359, 1991  
A:Title: Frog cardiac calsequestrin. Identification, characterization, and subcellula  
rdium.  
A:Reference number: A61230; MUID:91316784; PMID:1860177  
A:Accession: A61230  
A:Molecule type: protein  
A:Residues: 1-9 <MCL>  
C:Comment: Calsequestrin is a high-capacity and moderate-affinity calcium binding pro  
C:Comment: Calsequestrin acts as a calcium buffer, and the release of calcium bound t  
C:Comment: The fast skeletal muscle isoform of calsequestrin can be phosphorylated in  
C:Superfamily: calsequestrin  
C:Keywords: calcium binding; cardiac muscle; glycoprotein; heart; phosphoprotein; ske

Query Match 33.3%; Score 15; DB 2; Length 9;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 QEALAF 9

```
Db      1  EEGLNF 6
      :| | |
      1  EEGLNF 6

RESULT 3
S70721
heat shock protein C62.5 homolog - Salmonella typhimurium (fragment)
N:Alternate names: high temperature protein G
C:Species: Salmonella typhimurium
C:Date: 11-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 17-Apr-1998
C:Accession: S70721
R:Qi, S.Y.; Li, Y.; Szyroki, A.; Giles, I.G.; Moir, A.; O'Connor, C.D.
Mol. Microbiol. 17, 523-531, 1995
A:Title: Salmonella typhimurium responses to a bactericidal protein from human neutrophil
A:Reference number: S70719; MUID:96100451; PMID:8559071
A:Accession: S70721
A:Molecule type: protein
A:Residues: 1-10 <QIS>
A:Experimental source: strain SLL344
C:Keywords: ATP binding; heat shock; molecular chaperone

Query Match      33.38; Score 15; DB 2; Length 10;
Best Local Similarity 33.38; Pred. No. 5.2e+03;
Matches 3; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy      1  LMAQEALAF 9
      :| | |
      1  MKGQETRGF 9

Db      1  MKGQETRGF 9

RESULT 4
A37765
hypothetical protein (csma 5' region) - Chloroflexus aurantiacus (fragment)
C:Species: Chloroflexus aurantiacus
C:Date: 31-May-1991 #sequence_revision 31-May-1991 #text_change 30-Sep-1993
C:Accession: A37765
R:Theroux, S.J.; Redlinger, T.E.; Fuller, R.C.; Robinson, S.J.
J. Bacteriol. 172, 4497-4504, 1990
A:Title: Gene encoding the 5.7-kilodalton chlorosome protein of Chloroflexus aurantiacus
A:Reference number: A37765; MUID:90330558; PMID:2376566
A:Accession: A37765
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-6 <THE>
A:Cross-references: GB:M33964

Query Match      31.1%; Score 14; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2  MAQ 4
      :| |
      3  MAQ 5

Db      3  MAQ 5

RESULT 5
S66296
Na+-transporting ATP synthase (EC 3.6.1.-) chain c - Acetobacterium woodii (fragment)
N:Alternate names: ATPase chain c
C:Species: Acetobacterium woodii
C:Date: 19-Mar-1997 #sequence_revision 06-Jun-1997 #text_change 07-May-1999
C:Accession: S66296
R:Reidlinger, J.; Mueller, V.
Eur. J. Biochem. 223, 275-283, 1994
A:Title: Purification of ATP synthase from Acetobacterium woodii and identification as a
A:Reference number: S45648; MUID:94307271; PMID:8033902
A:Accession: S66296
A:Molecule type: protein
A:Residues: 1-8 <REI>
A:Experimental source: DSM 1030
C:Keywords: hydrolase

Query Match      31.1%; Score 14; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2  MAQ 4
      :| |
      3  MAQ 5

Db      3  MAQ 5

RESULT 6
F28854
fibrinopeptide B - olive baboon
C:Species: Papio anubis, Papio hamadryas anubis (olive baboon)
C:Date: 19-May-1989 #sequence_revision 19-May-1989 #text_change 18-Aug-2000
C:Accession: F28854
R:Nakamura, S.; Takenaka, O.; Takahashi, K.
J. Biochem. 94, 1973-1978, 1983
A:Title: Fibrinopeptides A and B of baboons (Papio anubis, Papio hamadryas, and Thero
A:Reference number: A91973; MUID:84161822; PMID:6423621
A:Accession: F28854
A:Molecule type: protein
A:Residues: 1-9 <NAK>
C:Superfamily: fibrinogen beta chain; fibrinogen beta/gamma homology; fibrinogen disu

Query Match      31.1%; Score 14; DB 2; Length 9;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      4  QEAL 7
      :| |
      2  QEGL 5

Db      2  QEGL 5

RESULT 7
E28854
fibrinopeptide B - hamadryas baboon
C:Species: Papio hamadryas (hamadryas baboon)
C:Date: 19-May-1989 #sequence_revision 19-May-1989 #text_change 18-Aug-2000
C:Accession: E28854
R:Nakamura, S.; Takenaka, O.; Takahashi, K.
J. Biochem. 94, 1973-1978, 1983
A:Title: Fibrinopeptides A and B of baboons (Papio anubis, Papio hamadryas, and Thero
A:Reference number: A91973; MUID:84161822; PMID:6423621
A:Accession: E28854
A:Molecule type: protein
A:Residues: 1-9 <NAK>
C:Superfamily: fibrinogen beta chain; fibrinogen beta/gamma homology; fibrinogen disu

Query Match      31.1%; Score 14; DB 2; Length 9;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      4  QEAL 7
      :| |
      2  QEGL 5

Db      2  QEGL 5

RESULT 8
F28854
fibrinopeptide B - gelada baboon
C:Species: Theropithecus gelada (gelada baboon)
C:Date: 19-May-1989 #sequence_revision 19-May-1989 #text_change 18-Aug-2000
C:Accession: F28854
R:Nakamura, S.; Takenaka, O.; Takahashi, K.
J. Biochem. 94, 1973-1978, 1983
A:Title: Fibrinopeptides A and B of baboons (Papio anubis, Papio hamadryas, and Thero
A:Reference number: A91973; MUID:84161822; PMID:6423621
A:Accession: F28854
A:Molecule type: protein
A:Residues: 1-9 <NAK>
C:Superfamily: fibrinogen beta chain; fibrinogen beta/gamma homology; fibrinogen disu

Query Match      31.1%; Score 14; DB 2; Length 9;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      4  QEAL 7
      :| |
      2  QEGL 5

Db      2  QEGL 5
```

Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 4 QEAL 7  
|||  
Db 2 QEGL 5

## RESULT 9

JN0027  
[Phe-6]-mosact - sea urchin (Clypeaster japonicus)  
C:Species: Clypeaster japonicus  
C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 20-Dec-1992  
C:Accession: JN0027  
R:Suzuki, N.; Kurita, M.; Yoshino, K.I.; Kajiura, H.; Nomura, K.; Yamaguchi, M.  
Zool. Sci. 4, 649-656, 1987  
A:Title: Purification and structure of mosact and its derivatives from the egg jelly of  
A:Reference number: JN0025  
A:Accession: JN0027  
A:Molecule type: protein  
A:Residues: 1-9 <SUZ>  
C:Comment: Mosact, one of several sperm activating peptides located in egg jelly, and it

Query Match 31.1%; Score 14; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 AFL 10  
|||  
Db 5 AFL 7

## RESULT 10

S24190  
tryptase (EC 3.4.21.59) - bovine (fragment)  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 19-Mar-1997 #sequence\_revision 18-Aug-2000 #text\_change 18-Aug-2000  
C:Accession: S24190  
R:Florucci, L.; Erba, F.; Ascoli, F.  
Biol. Chem. Hoppe-Seyler 373, 483-490, 1992  
A:Title: Bovine tryptase: purification and characterization.  
A:Reference number: S24190; MUID:92384956; PMID:1515079  
A:Accession: S24190  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-10 <PIO>  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; serine proteinase; zymogen

Query Match 31.1%; Score 14; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 8.5e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 QEA 6  
|||  
Db 5 QEA 7

## RESULT 11

A60722  
Cryptic fibrial protein - Serratia marcescens (strain US5) (fragment)  
C:Species: Serratia marcescens  
C:Date: 28-Apr-1993 #sequence\_revision 28-Apr-1993 #text\_change 17-Mar-1999  
C:Accession: A60722  
R:Moriya, T.; Kawabata, S.I.; Mizunoe, Y.; Amako, K.  
J. Bacteriol. 171, 6629-6636, 1989  
A:Title: A cryptic fibrial gene in Serratia marcescens.  
A:Reference number: A60722; MUID:90078108; PMID:2574174  
A:Accession: A60722  
A:Molecule type: protein  
A:Residues: 1-10 <MOR>  
C:Comment: This protein was purified after expression of its gene in E. coli. and appear

Query Match 31.1%; Score 14; DB 2; Length 10;

Best Local Similarity 75.0%; Pred. No. 8.5e+03;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 EALA 8  
|||  
Db 6 QALA 9

## RESULT 12

PH0113  
alpha-amylase (EC 3.2.1.1) III - rice (fragment)  
C:Species: Oryza sativa (rice)  
C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 07-May-1999  
C:Accession: PH0113  
R:Chiba, Y.; Nisida, Y.; Nakajima, T.; Ichishima, E.  
Agric. Biol. Chem. 55, 901-902, 1991  
A:Title: Unique enzymatic properties of alpha-amylase-III from suspension-cultured ri  
A:Reference number: PH0113; MUID:91234351; PMID:1368633  
A:Accession: PH0113  
A:Molecule type: protein  
A:Residues: 1-10 <CHI>  
A:Experimental source: cv. Sasanishiki  
C:Function:  
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds  
A:Pathway: glycogen/starch degradation  
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 31.1%; Score 14; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 8.5e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AOE 5  
|||  
Db 5 AOE 7

## RESULT 13

PT0246  
Ig heavy chain CRD3 region (clone 2-103D) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996  
C:Accession: PT0246  
R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.  
J. Exp. Med. 173, 395-407, 1991  
A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity an  
A:Reference number: PT0222; MUID:91108337; PMID:1899102  
A:Accession: PT0246  
A:Molecule type: DNA  
A:Residues: 1-7 <YAM>  
A:Experimental source: B lymphocyte  
C:Keywords: heterotetramer; immunoglobulin

Query Match 28.9%; Score 13; DB 2; Length 7;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 EALAF 10  
|||  
Db 2 EVLGIL 7

## RESULT 14

S66646  
cardioacceleratory protein 2b - tobacco hornworm  
C:Species: Manduca sexta (tobacco hornworm)  
C:Date: 15-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 17-Mar-1999  
C:Accession: S66646  
R:Huesmann, G.R.; Cheung, C.C.; Loi, P.K.; Lee, T.D.; Swiderek, K.M.; Tublitz, N.J.  
FEBS Lett. 371, 311-314, 1995  
A:Title: Amino acid sequence of CAP(2b), an insect cardioacceleratory peptide from th  
A:Reference number: S66646; MUID:96013159; PMID:7556618  
A:Accession: S66646  
A>Status: preliminary

A:Molecule type: protein  
A:Residues: 1-8 <HUE>

Query Match 28.9%; Score 13; DB 2; Length 8;  
Best Local Similarity 60.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 EALAF 9  
| | |  
Db 1 ELYAF 5

## RESULT 15

S65381  
cytochrome-c oxidase (EC 1.9.3.1) chain V1b, hepatic - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 07-May-1999  
C:Accession: S65381  
R:Schaegger, H.; Noack, H.; Halangk, W.; Brandt, U.; von Jagow, G.  
Eur. J. Biochem. 230, 235-241, 1995  
A:Title: Cytochrome-c oxidase in developing rat heart. Enzymic properties and amino-term  
A:Reference number: S65372; MUID:95324529; PMID:7601105  
A:Accession: S65381  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-8 <SCH>  
C:Keywords: oxidoreductase

Query Match 28.9%; Score 13; DB 2; Length 8;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 QEALAF 9  
| | | |  
Db 1 QNXLDF 6

Search completed: October 7, 2003, 14:31:16  
Job time : 25 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:17:20 ; Search time 13 Seconds  
(without alignments)  
36.174 Million cell updates/sec

Title: US-09-807-512-12  
Perfect score: 45  
Sequence: 1 LMAQALAF 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 372

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	18	40.0	7	1 ALL2_CARMA	P81805 carcinus ma
2	16	35.6	9	1 RT33_BOVIN	P82926 bos taurus
3	15	33.3	8	1 CLP_THICU	P80488 thioabaciliu
4	14	31.1	9	1 FIBB_PAPAN	P19344 papio anubi
5	14	31.1	9	1 FIBB_PAPHA	P19343 papio hamad
6	14	31.1	9	1 FIBB_THEGE	P19342 theropithec
7	14	31.1	9	1 MOSF_CLYJA	P19853 clypeaster
8	13	28.9	7	1 ALL3_CARMA	P81806 carcinus ma
9	13	28.9	8	1 COXG_RAT	P80430 rattus norv
10	13	28.9	9	1 ALI0_CARMA	P81813 carcinus ma
11	13	28.9	9	1 FAR2_CALVO	P41857 calliphora
12	13	28.9	9	1 FAR3_CALVO	P41858 calliphora
13	13	28.9	9	1 FRF1_SARBU	P83350 sarcophaga
14	13	28.9	9	1 ULAH_HUMAN	P31934 homo sapien
15	13	28.9	10	1 PSBF_CAPAN	P00367 capsicum an
16	13	28.9	10	1 URA6_HUMAN	P32080 homo sapien
17	12	26.7	7	1 FAR2_ASCSU	P31890 ascaris suu
18	12	26.7	9	1 FAR1_CALVO	P41856 calliphora
19	12	26.7	9	1 FAR4_CALVO	P41859 calliphora
20	12	26.7	9	1 FAR5_CALVO	P41860 calliphora
21	12	26.7	9	1 FAR6_CALVO	P41861 calliphora
22	12	26.7	9	1 FAR7_CALVO	P41862 calliphora
23	12	26.7	9	1 FIBB_MACFU	P19345 macaca fusc
24	12	26.7	10	1 FARP_MYTED	P42560 mytilus edu
25	12	26.7	10	1 TRP5_LEUMA	P81737 leucophaea
26	11	24.4	6	1 ACPH_RABIT	P25154 oryctolagus
27	11	24.4	6	1 TRPI_PSEPU	P36414 pseudomonas
28	11	24.4	7	1 CCF1_ENTFA	P20104 enterococcu
29	11	24.4	7	1 UFO3_MOUSE	P38641 mus musculu
30	11	24.4	8	1 FAR8_CALVO	P41863 calliphora
31	11	24.4	8	1 NS3_MYCTU	P81152 mycobacteri
32	11	24.4	8	1 UPAL_HUMAN	P30087 homo sapien
33	11	24.4	9	1 FAR3_PENMO	P83318 penaeus mon

RESULT 1  
ALL2\_CARMA  
ID ALL2\_CARMA STANDARD; PRT; 7 AA.

AC P81805;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Carcinustatin 2.  
OS Carcinus maenas (Common shore crab) (Green crab).  
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
OC Eubranchyura; Fortunoidea; Portunidae; Carcinus.  
OX NCBI\_TaxID=6759;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;  
RX MEDLINE=98121193; PubMed=9461295;  
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P., Thorpe A.;  
RT "Isolation and identification of multiple neuropeptides of the allatostatin superfamily in the shore crab Carcinus maenas.";  
RL Eur. J. Biochem. 250:727-734(1997).  
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.  
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
KW Neuropeptide; Amidation; Multigene family.  
FT MOD\_RES 7 7 AMIDATION (POTENTIAL).  
SQ SEQUENCE 7 AA; 770 MW; 672879CDB5DDB70 CRC64;

Query Match 40.0%; Score 18; DB 1; Length 7;  
Best Local Similarity 80.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 EALAF 9  
DB 1 EAYAF 5

RESULT 2  
RT33\_BOVIN  
ID RT33\_BOVIN STANDARD; PRT; 9 AA.  
AC P82926;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Mitochondrial 28S ribosomal protein S33 (S33mt) (MRP-S33) (Fragment).  
GN MRP33.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Liver;  
RX MEDLINE=21276436; PubMed=11279123;  
RA Koc E.C., Burkhardt W., Blackburn K., Moseley A., Spremulli L.L.;

34 11 24.4 9 1 FARP\_CALSI P38495 callinectes  
35 11 24.4 9 1 FIBB\_ERYPA P19346 erythrocebu  
36 11 24.4 9 1 HUTU\_KLEAE P12381 klebsiella  
37 11 24.4 9 1 SAMP\_MUSCA P19095 mustelus ca  
38 11 24.4 10 1 ESL\_LACCA P81758 lactobacill  
39 11 24.4 10 1 FARP\_MANSE P18523 manduca sex  
40 11 24.4 10 1 NO40\_TOBAC P55962 nicotiana t  
41 11 24.4 10 1 Q2OB\_COMTE P80465 comamonas t  
42 11 24.4 10 1 SPI\_HALRO Q10997 halocynthia  
43 11 24.4 10 1 TKLA\_LOCFMI P30250 locusta mig  
44 11 24.4 10 1 URAL\_HUMAN P32118 homo sapien  
45 10 22.2 4 1 DCMS\_PSECH P19918 pseudomonas

ALIGNMENTS

RT "The small subunit of the mammalian mitochondrial ribosome:  
 RT identification of the full complement of ribosomal proteins present.";

RL J. Biol. Chem. 276:19363-19374(2001).  
 CC -1- SUBUNIT: Component of the mitochondrial ribosome small subunit  
 CC (28S) which comprises a 12S rRNA and about 30 distinct proteins.  
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.  
 CC Ribosomal protein; Mitochondrion.  
 KW NON\_TER 1 1  
 FT NON\_TER 9 9  
 FT NON\_TER 9 9  
 SQ SEQUENCE 9 AA; 1032 MW; D341D73776DB05B9 CRC64;

Query Match 35.6%; Score 16; DB 1; Length 9;  
 Best Local Similarity 37.5%; Pred. No. 1.3e+05;  
 Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LMAQEALA 8  
 | : : : |  
 Db 1 LFSEQPLA 8

## RESULT 3

CLP\_THICU STANDARD; PRT; 8 AA.  
 AC P80488;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Chemolithotroph-specific protein (Fragment).  
 OS Thiobacillus cuprinus.  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Comamonadaceae; Thiomonas.  
 OX NCBI\_TaxID=36860;  
 [1]  
 RP SEQUENCE.  
 RC STRAIN=DSM 5494;  
 RA Marin I., Amaro A.M., Jerez C.A., Anils R., Abad J.P.;  
 RL Submitted (SEP-1995) to the SWISS-PROT data bank.  
 CC -1- MISCELLANEOUS: FOUND SPECIFICALLY IN CELLS CULTURED  
 CC CHEMOLITHOTROPHICALLY.  
 FT NON\_TER 8 8  
 SQ SEQUENCE 8 AA; 785 MW; 91487B06DDC2D76D CRC64;

Query Match 33.3%; Score 15; DB 1; Length 8;  
 Best Local Similarity 37.0%; Pred. No. 1.3e+05;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MAQE 5  
 : : : :  
 Db 3 VAQE 6

## RESULT 4

FIBB\_PAPAN STANDARD; PRT; 9 AA.  
 AC P19344;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).  
 GN FGB.  
 OS Papio anubis (Olive baboon).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;  
 OC Cercopitheciinae; Papio.  
 OX NCBI\_TaxID=9555;  
 [1]  
 RP SEQUENCE.  
 RX MEDLINE=84161822; PubMed=6423621;  
 RA Nakamura S., Takenaka O., Takahashi K.;  
 RT "Fibrinopeptides A and B of baboons (Papio anubis, Papio hamadryas,  
 RT and Theropithecus gelada): their amino acid sequences and  
 RT evolutionary rates and a molecular phylogeny for the baboons.";  
 RL J. Biochem. 94:1973-1978(1983).

CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
 CC AGGREGATION.  
 CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
 CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
 CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.

DR FIBRINAG\_C; D28854.  
 DR InterPro: IPR002181; Fibrinogen\_C.  
 DR PROSITE: PS00514; FIBRINAG\_C\_DOMAIN; PARTIAL.  
 KW Blood coagulation; Plasma.  
 FT PEPTIDE 1 9 FIBRINOPEPTIDE B.  
 FT NON\_TER 9 9  
 SQ SEQUENCE 9 AA; 1076 MW; DDFE6409C7287B06 CRC64;

Query Match 31.1%; Score 14; DB 1; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 1.3e+05;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 QEAL 7  
 | : : |  
 Db 2 QEGL 5

## RESULT 5

FIBB\_PAPHA STANDARD; PRT; 9 AA.  
 AC P19343;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).  
 GN FGB.  
 OS Papio hamadryas (Hamadryas baboon).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;  
 OC Cercopitheciinae; Papio.  
 OX NCBI\_TaxID=9557;  
 [1]  
 RP SEQUENCE.  
 RX MEDLINE=84161822; PubMed=6423621;  
 RA Nakamura S., Takenaka O., Takahashi K.;  
 RT "Fibrinopeptides A and B of baboons (Papio anubis, Papio hamadryas,  
 RT and Theropithecus gelada): their amino acid sequences and  
 RT evolutionary rates and a molecular phylogeny for the baboons.";  
 RL J. Biochem. 94:1973-1978(1983).  
 CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
 CC AGGREGATION.

CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
 CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
 CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.

DR FIBRINAG\_C; E28854.  
 DR InterPro: IPR002181; Fibrinogen\_C.  
 DR PROSITE: PS00514; FIBRINAG\_C\_DOMAIN; PARTIAL.  
 KW Blood coagulation; Plasma.  
 FT PEPTIDE 1 9 FIBRINOPEPTIDE B.  
 FT NON\_TER 9 9  
 SQ SEQUENCE 9 AA; 1057 MW; DDFE71E9C7287B06 CRC64;

Query Match 31.1%; Score 14; DB 1; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 1.3e+05;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 QEAL 7  
 | : : |  
 Db 2 QEGL 5

RESULT 6  
FIBB\_THEGE  
ID FIBB\_THEGE STANDARD; PRT; 9 AA.  
AC P19342;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).  
GN FGB.  
OS Theropithecus gelada (Gelada baboon).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecinae; Theropithecus.  
OX NCBI\_TaxID=9565;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=84161822; PubMed=6423621;  
RA Nakamura S., Takenaka O., Takahashi K.;  
RT "Fibrinopeptides A and B of baboons (Papio anubis, Papio hamadryas,  
RT and Theropithecus gelada): their amino acid sequences and  
RT evolutionary rates and a molecular phylogeny for the baboons.";  
RL J. Biochem. 94:1973-1978(1983).  
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
CC AGGREGATION.  
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.  
CC PIR; F28854; F28854.  
DR InterPro: IPR002181; Fibrinogen.C.  
DR PROSITE: PS00514; FIBRIN\_AG\_C\_DOMAIN; PARTIAL.  
KW Blood coagulation; Plasma.  
FT PEPTIDE 1 9 FIBRINOPEPTIDE B.  
FT NON\_TER 9 9  
SQ SEQUENCE 9 AA; 977 MW; DDFE7879C7287B06 CRC64;  
  
Query Match 31.1%; Score 14; DB 1; Length 9;  
Best Local Similarity 75.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 4 QEAL 7  
Db 2 QEGL 5  
  
RESULT 7  
MOSE\_CLYJA  
ID MOSE\_CLYJA STANDARD; PRT; 9 AA.  
AC P19853;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 01-FEB-1991 (Rel. 17, Last annotation update)  
DE [phe-6]-mosact.  
OS Clypeaster japonicus (Sand dollar).  
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
OC Echinoidea; Euechinoidea; Gnathostomata; Clypeasteroidea;  
OC Clypeasteridae; Clypeaster.  
OX NCBI\_TaxID=7644;  
RN [1]  
RP SEQUENCE.  
RX TISSUE=Egg jelly;  
RA Suzuki N., Kurita M., Yoshino K.I., Kajlura H., Nomura K.,  
RA Imaguchi M.;  
RT "Purification and structure of mosact and its derivatives from the  
RT egg jelly of the sea urchin Clypeaster japonicus";  
RL Zool. Sci. 4:649-656(1987).  
CC -!- FUNCTION: Stimulates sperm respiration and motility.  
DR PIR; JN0027; JN0027.  
SQ SEQUENCE 9 AA; 924 MW; 93245729CDC5B8B5 CRC64;

Query Match 31.1%; Score 14; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 8 AFL 10  
Db 5 AFL 7  
  
RESULT 8  
ALL3\_CARMA  
ID ALL3\_CARMA STANDARD; PRT; 7 AA.  
AC P81806;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Carcinustatin 3.  
OS Carcinus maenas (Common shore crab) (Green crab).  
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.  
OX NCBI\_TaxID=6759;  
RN [1]  
RP SEQUENCE.  
RX TISSUE=Cerebral ganglion, and Thoracic ganglion;  
RX MEDLINE=98121193; PubMed=9461295;  
RA Duve H., Johnsen A.H., Maestro J.L., Scott A.G., Jaros P.P.,  
RA Thorpe A.;  
RT "Isolation and identification of multiple neuropeptides of the  
RT allatostatin superfamily in the shore crab Carcinus maenas";  
RL Eur. J. Biochem. 250:727-734(1997).  
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.  
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
KW Neuropeptide; Multigene family.  
SQ SEQUENCE 7 AA; 796 MW; 672879CDCB476B70 CRC64;  
  
Query Match 28.9%; Score 13; DB 1; Length 7;  
Best Local Similarity 60.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 5 EALAF 9  
Db 1 EPYAF 5  
  
RESULT 9  
COXG\_RAT  
ID COXG\_RAT STANDARD; PRT; 8 AA.  
AC P80430;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cytochrome c oxidase polypeptide V1b (EC 1.9.3.1) (AED) (Fragment).  
GN COX6B.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE.  
RX STRAIN=Wistar; TISSUE=Liver;  
RX MEDLINE=95324529; PubMed=7601105;  
RA Schaeffer H., Noack H., Halangk W., Brandt U., von Jagow G.;  
RT "Cytochrome c oxidase in developing rat heart. Enzymic properties and  
RT amino-terminal sequences suggest identity of the fetal heart and the  
RT adult liver isoform";  
RL Eur. J. Biochem. 230:235-241(1995).  
CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE  
CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN  
CC MITOCHONDRIAL ELECTRON TRANSPORT. THIS PROTEIN MAY BE ONE OF THE  
CC HEME-BINDING SUBUNITS OF THE OXIDASE.  
CC -!- CATALYTIC ACTIVITY: 4 ferrocyanochrome c + O(2) -> 4 ferricytochrome

```
CC C + 2 H(2)O.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VIB FAMILY.
PIR; S65381; S65381.
KW Oxidoreductase; Mitochondrion.
DL NON_TER 1
FT NON_TER 8
SQ SEQUENCE 8 AA; 1039 MW; 8101E9CAA73AE456 CRC64;

Query Match 28.9%; Score 13; DB 1; Length 8;
Best Local Similarity 50.0%; Pred. No. 1.3e+05;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 AQEALAF 9
Db 1 QNXLDF 6

RESULT 10
AL10_CARMA STANDARD; PRT; 9 AA.
ID FAR2_CALVO
AC P41857;
DT 01-NOV-1995 (Rel. 32, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 10.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
RT allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation; Multigene family.
FT MOD_RES 9
SQ SEQUENCE 9 AA; 963 MW; 372D79CDBC4776C7 CRC64;

Query Match 28.9%; Score 13; DB 1; Length 9;
Best Local Similarity 42.9%; Pred. No. 1.3e+05;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 AQEALAF 9
Db 1 APQPYAF 7

RESULT 11
FAR2_CALVO STANDARD; PRT; 9 AA.
ID FAR2_CALVO
AC P41857;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE CallifMRamide 2.
OS Calliphora vomitoria (Blue blowfly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
OC Calliphoridae; Calliphora.
OX NCBI_TaxID=27454;
RN [1]
RP SEQUENCE.
RC TISSUE=Thoracic ganglion;
RX MEDLINE=92196111; PubMed=1549595;
RA Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,
RA Rehfeld J.F., Thorpe A.;
RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2
RT neuropeptides (designated callifMRamides) from the blowfly
RT Calliphora vomitoria.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).
CC -1- FUNCTION: ABLE TO INDUCE FLUID SECRETION FROM THE ISOLATED
CC SALIVARY GLAND OF CALLIPHORA.
CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
DR PIR; C41978; C41978.
KW Neuropeptide; Amidation.
FT MOD_RES 9
SQ SEQUENCE 9 AA; 1114 MW; 2F0B0699CAB6C5A7 CRC64;

Query Match 28.9%; Score 13; DB 1; Length 9;
Best Local Similarity 28.6%; Pred. No. 1.3e+05;
Matches 2; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 AQEALAF 9
Db 3 SQDFMRF 9

RESULT 13
FRF1_SARBU STANDARD; PRT; 9 AA.
ID FRF1_SARBU
AC P83350;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Neb-FMRFamide 1.
OS Sarcophaga bullata (Grey flesh fly) (Neobellieria bullata).
OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
```

```
RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2
RT neuropeptides (designated callifMRamides) from the blowfly
RT Calliphora vomitoria.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).
CC -1- FUNCTION: ABLE TO INDUCE FLUID SECRETION FROM THE ISOLATED
CC SALIVARY GLAND OF CALLIPHORA.
CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
DR PIR; B41978; B41978.
KW Neuropeptide; Amidation.
FT MOD_RES 9
SQ SEQUENCE 9 AA; 1128 MW; 29D00699CAB6C5A7 CRC64;

Query Match 28.9%; Score 13; DB 1; Length 9;
Best Local Similarity 28.6%; Pred. No. 1.3e+05;
Matches 2; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 AQEALAF 9
Db 3 SQDFMRF 9

RESULT 12
FAR3_CALVO STANDARD; PRT; 9 AA.
ID FAR3_CALVO
AC P41858;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE CallifMRamide 3.
OS Calliphora vomitoria (Blue blowfly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
OC Calliphoridae; Calliphora.
OX NCBI_TaxID=27454;
RN [1]
RP SEQUENCE.
RC TISSUE=Thoracic ganglion;
RX MEDLINE=92196111; PubMed=1549595;
RA Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,
RA Rehfeld J.F., Thorpe A.;
RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2
RT neuropeptides (designated callifMRamides) from the blowfly
RT Calliphora vomitoria.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).
CC -1- FUNCTION: ABLE TO INDUCE FLUID SECRETION FROM THE ISOLATED
CC SALIVARY GLAND OF CALLIPHORA.
CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
DR PIR; C41978; C41978.
KW Neuropeptide; Amidation.
FT MOD_RES 9
SQ SEQUENCE 9 AA; 1114 MW; 2F0B0699CAB6C5A7 CRC64;

Query Match 28.9%; Score 13; DB 1; Length 9;
Best Local Similarity 28.6%; Pred. No. 1.3e+05;
Matches 2; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 AQEALAF 9
Db 3 SQDFMRF 9
```

```

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
OC Sarcophagidae; Sarcophaga.
OX NCBI_TaxID=7385;
RN SEQUENCE, AMIDATION, AND FUNCTION.
RP TISSUE=CNS;
RX MEDLINE=22342733; PubMed=12438685;
RA Meusen T., Mertens I., Clynen E., Baggerman G., Nichols R.,
RA Nachman R.J., Huybrechts R., De Loof A., Schoofs L.;
RT "Identification in Drosophila melanogaster of the invertebrate G
RT protein-coupled FMRFamide receptor.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:15363-15368(2002).
CC -!- FUNCTION: Has modulatory actions at skeletal neuromuscular
CC junctions.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide; Amidation; Pyrrolidone carboxylic acid.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 9 9 AMIDATION.
SQ SEQUENCE 9 AA; 1155 MW; 2D810699CAB6C5A7 CRC64;

Query Match 28.9%; Score 13; DB 1; Length 9;
Best Local Similarity 28.6%; Pred. No. 1.3e+05;
Matches 2; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 AQEALAF 9
Db 3 SQDFMRF 9

RESULT 14
ULAH_HUMAN
ID ULAH_HUMAN STANDARD; PRT; 9 AA.
AC P31934;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Unknown protein from 2D-page of liver tissue (Spot 94) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=94147969; PubMed=8313870;
RA Hughes G.J., Frutiger S., Paquet N., Pasquali C., Sanchez J.-C.,
RA Tissot J.-D., Balroch A., Appel R.D., Hochstrasser D.F.;
RT "Human liver protein map: update 1993.";
RL Electrophoresis 14:1216-1222(1993).
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 5.5, ITS MW IS: 12 kDa.
DR SWISS-2DPAGE; P31934; HUMAN.
FT NON_TER 1 1
FT VARIANT 3 3 A -> L.
FT NON_TER 9 9 /FTIQ=VAR_000001.
FT SEQUENCE 9 AA; 990 MW; F1CC7B172AB6DDCA CRC64;

Query Match 28.9%; Score 13; DB 1; Length 9;
Best Local Similarity 42.9%; Pred. No. 1.3e+05;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 AQEALAF 9
Db 3 AODLEGF 9

RESULT 15
PSBF_CAPAN
ID PSBF_CAPAN STANDARD; PRT; 10 AA.
AC Q03367;

```

```

DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Cytochrome b559 beta subunit (PSII reaction center subunit VI)
DE (Fragment).
DE PSBF.
GN Capsicum annuum (Bell pepper).
OS Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamids; Solanales; Solanaceae; Capsicum.
OX NCBI_TaxID=4072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Lamuyo; TISSUE=Fruit, and Leaf;
RX MEDLINE=93099270; PubMed=1463853;
RA Kuntz M., Camara B., Weil J.-H., Schantz R.;
RT "The psbL gene from bell pepper (Capsicum annuum): plastid RNA
RT editing also occurs in non-photosynthetic chromoplasts.";
RL Plant Mol. Biol. 20:1185-1188(1992).
CC -!- FUNCTION: This b-type cytochrome is tightly associated with the
CC reaction center of photosystem II and possibly is part of the
CC water-oxidation complex.
CC -!- SUBUNIT: Heterodimer of an alpha subunit and a beta subunit.
CC -!- SIMILARITY: Belongs to the psbE / psbF family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X65570; CAA46539.1; -
DR PIR; S28055; S28055.
DR HAMAP; MF_00643; -; 1.
DR InterPro; IPR006216; CYT_b559.
DR PROSITE; PS00537; CYTOCHROME_B559; PARTIAL.
KW Chloroplast; Photosystem II; Heme; Electron transport; Transmembrane.
FT NON_TER 1 1
FT TRANSMEM <1 5 BY SIMILARITY.
FT DOMAIN 6 10 LUMENAL (POTENTIAL).
FT SEQUENCE 10 AA; 1180 MW; 817D0F59D6D69DC5 CRC64;

Query Match 28.9%; Score 13; DB 1; Length 10;
Best Local Similarity 40.0%; Pred. No. 7e+03;
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 6 ALAFL 10
Db 4 AMQFI 8

Search completed: October 7, 2003, 14:28:16
Job time : 13 secs

```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:24:10 ; Search time 58.5 seconds  
(without alignments)  
44.112 Million cell updates/sec

Title: US-09-807-512-12  
Perfect score: 45  
Sequence: 1 LMAQALAF 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 1349

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_23.3\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	40.0	9	13 O8AYL5	O8ayl5 carassius a
2	18	40.0	10	2 Q52762	Q52762 rhizobium s
3	18	40.0	10	12 Q66190	Q66190 avian infec
4	17	37.8	8	2 P83152	P83152 anabaena sp
5	17	37.8	10	2 O8RIT1	O8rit1 anaplasma p
6	16	35.6	9	2 Q46179	Q46179 clostridium
7	15	33.3	9	2 P72149	P72149 pseudomonas
8	15	33.3	9	2 P83157	P83157 anabaena sp
9	15	33.3	9	4 Q9NYH5	Q9nyh5 homo sapien
10	15	33.3	9	5 Q9TFV1	Q9tfv1 trypanosoma
11	15	33.3	9	13 Q8AUM7	Q8aum7 carassius a
12	15	33.3	10	10 O8LRT7	O8lrt7 chlamydomon
13	15	33.3	10	11 O70580	O70580 mus musculus
14	14	31.1	9	6 Q8MJT7	Q8mjt7 eulemur ful
15	14	31.1	9	6 Q8MJT8	Q8mjt8 eulemur ful
16	14	31.1	9	10 Q9AXH8	Q9axh8 mesembryant

17	14	31.1	10	2 Q56097	Q56097 salmonella
18	14	31.1	10	2 Q9R7J8	Q9r7j8 helicobacte
19	14	31.1	10	6 Q9TQV4	Q9tav4 equus cabal
20	14	31.1	10	11 Q8CJEO	Q8cje0 rattus norv
21	14	31.1	10	11 Q8BHN2	Q8bhn2 mus musculu
22	14	31.1	10	12 Q9L1U0	Q9qlu0 salmirine
23	14	31.1	10	12 P90373	P90373 pseudorabie
24	14	31.1	10	13 O73588	O73588 gallus gall
25	13	28.9	8	15 Q98YK9	Q98yk9 human immun
26	13	28.9	9	4 Q9BX14	Q9bx14 homo sapien
27	13	28.9	9	10 P82440	P82440 nicotiana t
28	13	28.9	9	12 Q91BY9	Q91by9 kaposi's sa
29	13	28.9	9	15 Q85710	Q85710 rous sarcom
30	13	28.9	10	2 Q9AE19	Q9ae19 streptococc
31	13	28.9	10	5 Q95NT8	Q95nt8 musca domes
32	13	28.9	10	5 Q8WPL6	Q8wpl6 oikopleura
33	13	28.9	10	8 Q8WGD2	Q8wgd2 hepatus eph
34	13	28.9	10	10 Q41788	Q41788 zea mays (m
35	13	28.9	10	10 Q8L7F5	Q8l7f5 hevea brasi
36	13	28.9	10	11 Q9QVK8	Q9qvk8 mus sp. mep
37	13	28.9	10	12 Q9W8B6	Q9w8b6 hepatitis g
38	13	28.9	10	12 Q39948	Q39948 hepatitis g
39	13	28.9	10	12 Q90344	Q90344 hepatitis g
40	12	26.7	7	4 Q15903	Q15903 homo sapien
41	12	26.7	8	2 Q56759	Q56759 xanthobacte
42	12	26.7	8	2 O09258	O09258 synechococc
43	12	26.7	8	2 Q8FSR3	Q8fsr3 lactobacill
44	12	26.7	8	2 O52062	O52062 bacillus me
45	12	26.7	8	10 Q9SAY7	Q9say7 dioscorea t

ALIGNMENTS

RESULT 1

ID	Q8AYL5	PRELIMINARY;	PRT;	9 AA.
AC	Q8AYL5;			
DT	01-MAR-2003 (TREMBLrel. 23, Created)			
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DE	Cytochrome P450 aromatase (Fragment).			
GN	CYP19A.			
OS	Carassius auratus (Goldfish).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;			
OC	Cyprinidae; Carassius.			
OX	NCBI_TaxID=7957;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Tchoudakova A.V., Kishida M., Wood E., Callard G.V.;			
RT	"Promoter characteristics of two CYP19 genes differentially expressed in the brain and ovary of teleost fish."			
RL	J. Steroid Biochem. Mol. Biol. 0:0-0(2001).			
FR	EMBL; AF324897; AAN32618.1; -.			
FT	NON_TER			
SQ	SEQUENCE 9 AA; 1060 MW; C49E76D7272B040D CRC64;			

Query Match 40.0%; Score 18; DB 13; Length 9;  
Best Local Similarity 66.7%; Pred. No. 8.3e+05;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy	2 MAQEAL 7
Db	1 MARELL 6

RESULT 2

ID	Q52762	PRELIMINARY;	PRT;	10 AA.
AC	Q52762;			
DT	01-NOV-1996 (TREMBLrel. 01, Created)			
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)			

```
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Dinitrogenase beta-subunit (Fragment).
OS Rhizobium sp. cowpea (strain IRC78).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=400;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IRC78;
RA Yun A.C., Szalay A.A.;
RT "Structural genes of dinitrogenase and dinitrogenase reductase are
RT transcribed from two separate promoters in the broad host range cowpea
RT Rhizobium strain IRC78."
RL Proc. Natl. Acad. Sci. U.S.A. 81:7358-7362(1984).
DR EMBL; M10203; AAA26309.1; -.
FT NON_TER 10
SQ SEQUENCE 10 AA; 1100 MW; CFFD0BC1EB1DC5A6 CRC64;

Query Match 40.0%; Score 18; DB 2; Length 10;
Best Local Similarity 80.0%; Pred. No. 4.2e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 MAQEA 6
Db 1 MAQSA 5

RESULT 3
Q66190 PRELIMINARY; PRT; 10 AA.
AC Q66190;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 1.1 kDa protein (Fragment).
OS Avian infectious bronchitis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=11120;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=41;
RX MEDLINE=87021475; PubMed=2429473;
RA Niesters H.G., Lenstra J.A., Spaan W.J., Zijderveld A.J.,
RA Bleumink-Pluym N.M., Hong F., van Scharrenburg G.J., Horzinek M.C.,
RA van der Zeijst B.A.;
RT "The peplomer protein sequence of the M41 strain of coronavirus IBV
RT and its comparison with Beaudette strains."
RL Virus Res. 5:253-263(1986).
DR EMBL; M21883; AAA66576.1; -.
KW Hypothetical protein.
FT NON_TER 10
SQ SEQUENCE 10 AA; 1136 MW; CF2510D5A1B775A6 CRC64;

Query Match 40.0%; Score 18; DB 12; Length 10;
Best Local Similarity 44.4%; Pred. No. 4.2e+03;
Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 MAQEAALFL 10
Db 1 MIQSPTSFL 9

RESULT 4
P83152 PRELIMINARY; PRT; 8 AA.
AC P83152;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Allophycocyanin beta chain (Fragment).
OS Anabaena sp. (strain L31).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.
```

```
OX NCBI_TaxID=29412;
RN [1]
RP SEQUENCE.
RA Apte S.K., Uhlemann E., Schmid R., Altendorf K.;
RL Submitted (OCT-2001) to the SWISS-PROT data bank.
CC -!- FUNCTION: (OCT-2001) THE SWISS-PROT data bank.
CC FROM THE PHYCOBILIPROTEIN COMPLEX. ALLOPHYCOCYANIN HAS A MAXIMUM
CC ABSORPTION AT APPROXIMATELY 650 TO 653 NANOMETERS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC -!- SUBCELLULAR LOCATION: THYLAKOID MEMBRANE.
CC -!- PTM: CONTAINS ONE COVALENTLY LINKED BILIN CHROMOPHORE.
KW Phycobilisome; Electron transport; Photosynthesis; Bile pigment;
KW Thylakoid; Membrane.
FT NON_TER 8
SQ SEQUENCE 8 AA; 788 MW; 87CDC1A05DDAB6DD CRC64;

Query Match 37.8%; Score 17; DB 2; Length 8;
Best Local Similarity 60.0%; Pred. No. 8.3e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 AQEAL 7
Db 1 AQDAI 5

RESULT 5
Q8RIT1 PRELIMINARY; PRT; 10 AA.
ID Q8RIT1
AC Q8RIT1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Heat shock protein groES (Fragment).
GN GROES.
OS Anaplasma phagocytophilum (Ehrlichia phagocytophila).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Anaplasmataceae; Anaplasma.
OX NCBI_TaxID=948;
RN [1]
RP SEQUENCE FROM N.A.
RA von Loewenich F.D., Bogdan C.;
RT "A case of equine granulocytic ehrlichiosis provides further evidence
RT for the presence of the HGE agent in Germany."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF482760; AAL8676.1; -.
FT NON_TER 1
SQ SEQUENCE 10 AA; 1103 MW; 9792B43DD0505AB5 CRC64;

Query Match 37.8%; Score 17; DB 2; Length 10;
Best Local Similarity 50.0%; Pred. No. 7e+03;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LMAQEA 6
Db 5 IIAKEA 10

RESULT 6
Q46179 PRELIMINARY; PRT; 9 AA.
ID Q46179
AC Q46179;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Chloramphenicol acetyltransferase.
GN CATQ.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CW 531;
```



```

RX MEDLINE=91247774; PubMed=2039197;
RA Bannam T.L., Rood J.I.;
RT "The relationship between the Clostridium perfringens catQ gene
RT product and chloramphenicol acetyltransferases from other bacteria.";
RL Antimicrob. Agents Chemother. 35:471-476(1991).
DR EMBL; M55620; AAA23214.1; -.
SQ SEQUENCE 9 AA; 1041 MW; AFF472322CDD696 CRC64;

Query Match 35.6%; Score 16; DB 2; Length 9;
Best Local Similarity 44.4%; Pred. No. 8.3e+05;
Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LMAQEALAF 9
   :|||
Db 1 MMMAVKLAF 9

RESULT 7
P72149
ID P72149 PRELIMINARY; PRT; 9 AA.
AC P72149;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Putative glucokinase (Fragment).
GN GLK.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PA01.
RX MEDLINE=96427344; PubMed=8830708;
RA Sage A.E., Proctor W.D., Phibbs P.V.Jr.;
RT "A two-component response regulator, gltr, is required for glucose
RT transport activity in Pseudomonas aeruginosa PA01.";
RL J. Bacteriol. 178:6064-6066(1996).
DR EMBL; U50932; AAC44474.1; -.
KW Kinase.
FT NON_TER 1
SQ SEQUENCE 9 AA; 944 MW; C3071DDAA72DC6C6 CRC64;

Query Match 33.3%; Score 15; DB 2; Length 9;
Best Local Similarity 75.0%; Pred. No. 8.3e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 QEALAF 7
   :|||
Db 2 QOAL 5

RESULT 8
P83157
ID P83157 PRELIMINARY; PRT; 9 AA.
AC P83157;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Superoxide dismutase [Fe] (EC 1.15.1.1) (Fragment).
OS Anabaena sp. (strain L31).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.
OX NCBI_TaxID=29412;
RN [1]
RP SEQUENCE.
RA Apte S.K., Uhlemann E., Schmid R., Altendorf K.;
RL Submitted (OCT-2001) to the SWISS-PROT data bank.
CC -!- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS.
CC -!- CATALYTIC ACTIVITY: 2 PEROXIDE RADICAL + 2 H(+) = O(2) + H(2)O(2).
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
```

```

CC -!- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
CC FAMILY.
DR InterPro: IPR001189; SODismutase.
DR PROSITE: PS00088; SOD_MN; PARTIAL.
KW Oxidoreductase; Iron; Metal-binding.
FT NON_TER 9
SQ SEQUENCE 9 AA; 1063 MW; C54267376B06C2C9 CRC64;

Query Match 33.3%; Score 15; DB 2; Length 9;
Best Local Similarity 50.0%; Pred. No. 8.3e+05;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 QEALAF 9
   :|||
Db 4 QEPLPY 9

RESULT 9
Q9NYH5
ID Q9NYH5 PRELIMINARY; PRT; 9 AA.
AC Q9NYH5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein KIAA0020 (Fragment).
GN KIAA0020.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Akatsuka Y., Brickner A.G., Warren E.H., Engelhard V.H., Riddell S.R.;
RT "Determination of intronic sequences adjacent to exonic sequence using
RT PCR and a genomic DNA library constructed by TA cloning.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF245442; AAF69838.1; -.
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 9
SQ SEQUENCE 9 AA; 1029 MW; 0A73DDD68771A720 CRC64;

Query Match 33.3%; Score 15; DB 4; Length 9;
Best Local Similarity 75.0%; Pred. No. 8.3e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MAQE 5
   :|||
Db 6 MAQK 9

RESULT 10
Q9TVF1
ID Q9TVF1 PRELIMINARY; PRT; 9 AA.
AC Q9TVF1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE Mucin-like protein (Fragment).
GN EMUCE-19C8.
OS Trypanosoma cruzi.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5693;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C1-Brenner;
RX MEDLINE=98225151; PubMed=9556557;
RA Di Noia J.M., D'Orso I., Aslund L., Sanchez D.O., Frasch A.C.;
RT "The Trypanosoma cruzi mucin family is transcribed from hundreds of
RT genes having hypervariable regions.";
RL J. Biol. Chem. 273:10843-10850(1998).
DR EMBL; AF036447; AAC14246.1; -.
FT NON_TER 1
```

SQ SEQUENCE 9 AA; 896 MW; DBA831B1BB5DD72D CRC64;

Query Match 33.3%; Score 15; DB 5; Length 9;  
Best Local Similarity 75.0%; Pred. No. 8.3e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 ALAF 9  
|||  
2 ALAY 5

RESULT 11

Q8AUM7 PRELIMINARY; PRT; 9 AA.

ID O8AUM7  
AC O8AUM7  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Cytochrome P450 aromatase (Fragment).  
GN CYP19A.  
OS Carassius auratus (Goldfish).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Carassius.  
OX NCBI\_TaxID=7957;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Tchoudakova A.V., Kishida M., Wood E., Callard G.V.;  
RT "Promoter characteristics of two CYP19 genes differentially expressed  
in the brain and ovary of teleost fish.";  
RL J. Steroid Biochem. Mol. Biol. 0:0-0(2001).  
DR EMBL; AF324895; AAN32616.1; -.  
DR EMBL; AF324896; AAN32617.1; -.  
FT NON\_TER  
SQ SEQUENCE 9 AA; 961 MW; C49E76D7272B187D CRC64;

Query Match 33.3%; Score 15; DB 13; Length 9;  
Best Local Similarity 66.7%; Pred. No. 8.3e+05;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 MAQEL 7  
|||  
1 MAGELL 6

RESULT 12

Q8LRT7 PRELIMINARY; PRT; 10 AA.

ID Q8LRT7  
AC Q8LRT7  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE zslb (Fragment).  
OS Chlamydomonas reinhardtii.  
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
OC Chlamydomonadaceae; Chlamydomonas.  
OX NCBI\_TaxID=3055;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CC-2290;  
RA Kathir P., Lavoie M., Lefebvre P.A., Silflow C.D.;  
RT "Molecular Map of the Chlamydomonas Nuclear Genome.";  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF503638; AAM27452.1; -.  
FT NON\_TER  
SQ SEQUENCE 10 AA; 1071 MW; C7D8DC6C6DC6C6 CRC64;

Query Match 33.3%; Score 15; DB 10; Length 10;  
Best Local Similarity 75.0%; Pred. No. 1.9e+04;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 AQEA 6  
|||

Db 5 AQQA 8

RESULT 13

Q7O580 PRELIMINARY; PRT; 10 AA.

ID Q7O580  
AC Q7O580  
DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Carbonic anhydrase III (Fragment).  
GN CAR3 OR CAR5 OR CAR3.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Sowden J., Smith H., Morrison K., Edwards Y.;  
RT "Sequence comparisons and functional studies of the proximal promoter  
of the carbonic anhydrase 3 (CA3) gene.";  
RL Gene 214:157-165(1999).  
DR EMBL; AJ006474; CAA07057.1; -.  
DR MGD; MGI:88270; Car3.  
DR MGD; MGI:1351477; Car3.  
FT NON\_TER  
SQ SEQUENCE 10 AA; 1105 MW; 56F7FE71ADC37B13 CRC64;

Query Match 33.3%; Score 15; DB 11; Length 10;  
Best Local Similarity 75.0%; Pred. No. 1.9e+04;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MAOE 5  
|||  
1 MAKE 4

RESULT 14

Q8MJT7 PRELIMINARY; PRT; 9 AA.

ID Q8MJT7  
AC Q8MJT7  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Hemopexin (Fragment).  
OS Eulemur fulvus (brown lemur).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Strepsirhini; Lemnidae; Eulemur.  
OX NCBI\_TaxID=13515;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Wyner Y.M., Johnson S.E., Stumpf R., DeSalle R.;  
RT "A genetic assessment of a red-fronted/white-collared lemur hybrid  
zone at Andringitra, Madagascar.";  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF258049; AAM43870.1; -.  
FT NON\_TER  
SQ SEQUENCE 9 AA; 1110 MW; 738CC9C720587B1B CRC64;

Query Match 31.1%; Score 14; DB 6; Length 9;  
Best Local Similarity 33.3%; Pred. No. 8.3e+05;  
Matches 2; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 QEALAF 9  
:|:|  
2 EEGILF 7

RESULT 15

Q8MJT8 PRELIMINARY; PRT; 9 AA.

ID Q8MJT8

AC Q8MJT8;

DT 01-OCT-2002 (TREMBLrel. 22, Created)  
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)  
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
DE Hemopexin (Fragment).  
OS Eulemur fulvus albocollaris.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Strepsirhini; Lemnidae; Eulemur.  
OX NCBI\_TaxID=122224;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Wyner Y.M., Johnson S.E., Stumpf R., Desalle R.;  
RT "A genetic assessment of a red-fronted/white-collared lemur hybrid  
zone at Andringitra, Madagascar."  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF258028; AAM43849.1; -.  
FT NON\_TER 1  
FT NON\_TER 9  
SQ SEQUENCE 9 AA; 1110 MW; 738CC9C720587B1B CRC64;  
  
Query Match 31.1%; Score 14; DB 6; Length 9;  
Best Local Similarity 33.3%; Pred. No. 8.3e+05;  
Matches 2; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
  
QY 4 QALAF 9  
: | : |  
Db 2 EGGILF 7  
  
Search completed: October 7, 2003, 14:30:22  
Job time : 60.5 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:25:50 ; Search time 19.5 seconds  
(without alignments)  
21.698 Million cell updates/sec

Title: US-09-807-512-12

Perfect score: 45

Sequence: 1 LMAQETALAF 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 90058

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*

2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*

3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*

4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*

5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*

6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	24	53.3	9	1	US-08-036-555B-27
2	24	53.3	9	1	US-08-469-569-27
3	24	53.3	9	1	US-08-249-322A-27
4	24	53.3	9	1	US-08-469-526A-27
5	24	53.3	9	2	US-08-734-591A-27
6	24	53.3	9	2	US-08-469-660-27
7	24	53.3	9	3	US-08-470-335-27
8	24	53.3	9	3	US-08-735-021-27
9	24	53.3	9	3	US-08-734-664A-27
10	24	53.3	9	3	US-08-470-339-27
11	24	53.3	9	4	US-08-467-602-27
12	24	53.3	9	5	PCT-US94-05083C-27
13	24	53.3	9	5	PCT-US95-06846A-27
14	24	53.3	10	1	US-08-036-555B-14
15	24	53.3	10	1	US-08-469-569-14
16	24	53.3	10	1	US-08-249-322A-14
17	24	53.3	10	1	US-08-469-526A-14
18	24	53.3	10	2	US-08-734-591A-14
19	24	53.3	10	2	US-08-469-660-14
20	24	53.3	10	3	US-08-470-335-14
21	24	53.3	10	3	US-08-735-021-14
22	24	53.3	10	3	US-08-734-664A-14
23	24	53.3	10	3	US-08-470-339-14
24	24	53.3	10	4	US-08-467-602-14
25	24	53.3	10	5	PCT-US94-05083C-14
26	24	53.3	10	5	PCT-US95-06846A-14
27	24	53.3	10	5	PCT-US95-06846A-14

Sequence 28, Appl  
Sequence 29, Appl  
Sequence 30, Appl  
Sequence 4, Appl  
Sequence 31, Appl  
Sequence 9, Appl  
Sequence 34, Appl  
Sequence 27, Appl  
Sequence 27, Appl  
Sequence 27, Appl  
Sequence 14, Appl  
Sequence 61, Appl  
Sequence 61, Appl  
Sequence 73, Appl  
Sequence 309, Appl  
Sequence 8, Appl

28 22 48.9 9 4 US-09-434-476A-28  
29 22 48.9 9 4 US-09-434-476A-29  
30 22 48.9 9 4 US-09-434-476A-30  
31 21 46.7 9 2 US-08-408-095-4  
32 21 46.7 9 4 US-09-434-476A-31  
33 21 46.7 10 4 US-09-604-957-9  
34 20 44.4 6 1 US-08-260-202A-34  
35 20 44.4 8 2 US-08-679-865-27  
36 20 44.4 8 2 US-08-680-876-27  
37 20 44.4 8 3 US-09-263-975-27  
38 20 44.4 8 4 US-09-129-192C-27  
39 20 44.4 9 2 US-08-370-909-14  
40 20 44.4 9 2 US-08-637-759B-61  
41 20 44.4 9 3 US-08-871-355A-61  
42 20 44.4 9 4 US-09-201-945-61  
43 20 44.4 9 4 US-09-644-600-73  
44 20 44.4 9 4 US-09-289-198-309  
45 19 42.2 8 1 US-08-178-476A-8

#### ALIGNMENTS

#### RESULT 1

US-08-036-555B-27

; Sequence 27, Application US/08036555B

; Patent No. 5530109

; GENERAL INFORMATION:

; APPLICANT: Goodheartl, Andrew; Stroobant, Paul;

; APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;

; APPLICANT: Chen, Maio Su; Hiles, Ian

; TITLE OF INVENTION: Glial Mitogenic Factors, Their

; TITLE OF INVENTION: Preparation and Use

; NUMBER OF SEQUENCES: 184

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Felfe & Lynch.

; STREET: 805 Third Avenue

; CITY: New York City

; STATE: New York

; COUNTRY: USA

; ZIP: 10022

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

; COMPUTER: IBM

; OPERATING SYSTEM: PC-DOS

; SOFTWARE: Wordperfect

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/036,555B

; FILING DATE: 24-MAR-1993

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/965,173

; FILING DATE: 23-OCT-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/940,389

; FILING DATE: 03-SEP-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/907,138

; FILING DATE: 30-JUN-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/863,703

; FILING DATE: 03-APRIL-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: U.K. 91 07566.3

; FILING DATE: 10-APRIL-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Tsai, Christine H.

; REGISTRATION NUMBER: 34,266

; REFERENCE/DOCKET NUMBER: LUD 5250.4

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 688-9200

; TELEFAX: (212) 838-3884

; INFORMATION FOR SEQ ID NO: 27:

## ; SEQUENCE CHARACTERISTICS:

; LENGTH: 9  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; US-08-036-555B-27

Query Match 53.3%; Score 24; DB 1; Length 9;  
Best Local Similarity 75.0%; Pred. No. 2.5e+05;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 AQEALAEFL 10  
Db 1 AKEALAAAL 8

## RESULT 2

US-08-469-569-27  
; Sequence 27, Application US/08469569  
; Patent No. 5606032

## ; GENERAL INFORMATION:

; APPLICANT: Goodearl, Andrew; Stroobant, Paul;  
; APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchionni, Mark;  
; APPLICANT: Chen, Maio Su; Hiles, Ian  
; TITLE OF INVENTION: Glial Mitogenic Factors, Their  
; TITLE OF INVENTION: Preparation and Use  
; NUMBER OF SEQUENCES: 184

## ; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Felife & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022

## ; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
; COMPUTER: IBM

## ; OPERATING SYSTEM: PC-DOS

## ; SOFTWARE: Wordperfect

## ; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/469,569

; FILING DATE: 06-JUN-1995

; CLASSIFICATION: 530

## ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/036,555

; FILING DATE: 24-MAR-1993

; APPLICATION NUMBER: 07/965,173

; FILING DATE: 23-OCT-1992

## ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/940,389

; FILING DATE: 03-SEP-1992

## ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/907,138

; FILING DATE: 30-JUN-1992

## ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/863,703

; FILING DATE: 03-APRIL-1992

## ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: U.K. 91 07566.3

; FILING DATE: 10-APRIL-1991

## ; ATTORNEY/AGENT INFORMATION:

; NAME: Tsai, Christine H.

; REGISTRATION NUMBER: 34,266

; REFERENCE/DOCKET NUMBER: LUD 5250.4

; TELEPHONE: (212) 688-9200

; TELEFAX: (212) 838-3884

; INFORMATION FOR SEQ ID NO: 27:

## ; SEQUENCE CHARACTERISTICS:

; LENGTH: 9

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

US-08-469-569-27

Query Match 53.3%; Score 24; DB 1; Length 9;  
Best Local Similarity 75.0%; Pred. No. 2.5e+05;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 AQEALAEFL 10  
Db 1 AKEALAAAL 8

## RESULT 3

US-08-249-322A-27  
; Sequence 27, Application US/08249322A  
; Patent No. 5716930

## ; GENERAL INFORMATION:

; APPLICANT: Goodearl, Andrew; Stroobant, Paul;  
; APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchionni, Mark;  
; APPLICANT: Chen, Maio Su; Hiles, Ian  
; TITLE OF INVENTION: Glial Mitogenic Factors, Their  
; TITLE OF INVENTION: Preparation and Use  
; NUMBER OF SEQUENCES: 184

## ; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Felife & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022

## ; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
; COMPUTER: IBM

## ; OPERATING SYSTEM: PC-DOS

## ; SOFTWARE: Wordperfect

## ; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/249,322A

; FILING DATE: 26-MAY-1994

; CLASSIFICATION: 435

## ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/036,555

; FILING DATE: 24-MAR-1993

## ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/965,173

; FILING DATE: 23-OCT-1992

## ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/940,389

; FILING DATE: 03-SEP-1992

## ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/907,138

; FILING DATE: 30-JUN-1992

## ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/863,703

; FILING DATE: 03-APRIL-1992

## ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: U.K. 91 07566.3

; FILING DATE: 10-APRIL-1991

## ; ATTORNEY/AGENT INFORMATION:

; NAME: Tsai, Christine H.

; REGISTRATION NUMBER: 34,266

; REFERENCE/DOCKET NUMBER: LUD 250.4

; TELEPHONE: (212) 688-9200

; TELEFAX: (212) 838-3884

; INFORMATION FOR SEQ ID NO: 27:

## ; SEQUENCE CHARACTERISTICS:

; LENGTH: 9

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

US-08-249-322A-27

Query Match 53.3%; Score 24; DB 1; Length 9;  
Best Local Similarity 75.0%; Pred. No. 2.5e+05;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 AQEALAF 10  
I:|||||  
Db 1 AKEALAL 8

## RESULT 4

US-08-469-526A-27  
; Sequence 27, Application US/08469526A  
; Patent No. 5792849  
; GENERAL INFORMATION:  
; APPLICANT: Goodearl, Andrew  
; APPLICANT: Stroobant, Paul  
; APPLICANT: Minghetti, Luisa  
; APPLICANT: Waterfield, Michael  
; APPLICANT: Marchionni, Mark  
; APPLICANT: Chen, Maio Su  
; APPLICANT: Hiles, Ian  
; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR  
; TITLE OF INVENTION: PREPARATION AND USE  
; NUMBER OF SEQUENCES: 187  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Clark & Elbing LLP  
; STREET: 176 Federal Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/469,526A  
FILING DATE: 06 June 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/036,555  
FILING DATE: 24-MAR-1993  
APPLICATION NUMBER: 07/965,173  
FILING DATE: 23-OCT-1992  
APPLICATION NUMBER: 07/940,389  
FILING DATE: 03-SEP-1992  
APPLICATION NUMBER: 07/907,138  
FILING DATE: 03-JUN-1992  
APPLICATION NUMBER: 07/863,703  
FILING DATE: 03-APRIL-1992  
APPLICATION NUMBER: U.K. 91 07566.3  
FILING DATE: 10-APR-1991

ATTORNEY/AGENT INFORMATION:  
NAME: Bieker-Brady, Kristina  
REGISTRATION NUMBER: 39,109  
REFERENCE/DOCKET NUMBER: 04585/00200A  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

US-08-469-526A-27

Query Match 53.3%; Score 24; DB 1; Length 9;  
Best Local Similarity 75.0%; Pred. No. 2.5e+05;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 AQEALAF 10  
I:|||||  
Db 1 AKEALAL 8

## RESULT 5

US-08-734-591A-27  
; Sequence 27, Application US/08734591A  
; Patent No. 5854220  
; GENERAL INFORMATION:  
; APPLICANT: Goodearl, Andrew  
; APPLICANT: Stroobant, Paul  
; APPLICANT: Minghetti, Luisa  
; APPLICANT: Waterfield, Michael  
; APPLICANT: Hiles, Ian  
; APPLICANT: Marchionni, Mark  
; APPLICANT: Chen, Mario  
; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR  
; TITLE OF INVENTION: PREPARATION AND USE  
; NUMBER OF SEQUENCES: 187  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Clark & Elbing LLP  
; STREET: 176 Federal Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 MB  
COMPUTER: IBM Compatible Pentium  
OPERATING SYSTEM: Windows95  
SOFTWARE: WordPerfect (Version 7.0)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/734,591A  
FILING DATE: 22-OCT-1996  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/470,335  
FILING DATE: 06-JUN-1995

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/036,555  
FILING DATE: 03-MAR-1993  
APPLICATION NUMBER: 07/965,173  
FILING DATE: 23-OCT-1992  
APPLICATION NUMBER: 07/940,389  
FILING DATE: 03-SEP-1992  
APPLICATION NUMBER: 07/907,138  
FILING DATE: 30-JUN-1992  
APPLICATION NUMBER: 07/863,703  
FILING DATE: 03-APR-1992  
APPLICATION NUMBER: UK 91 07566.3  
FILING DATE: 10-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Bieker-Brady, Kristina  
REGISTRATION NUMBER: 39,109  
REFERENCE/DOCKET NUMBER: 04585/00200P  
TELEPHONE: (617) 428-0200  
TELEFAX: (617) 428-7045  
TELEX:

INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear

US-08-734-591A-27

Query Match 53.3%; Score 24; DB 2; Length 9;  
Best Local Similarity 75.0%; Pred. No. 2.5e+05;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 AOEALAF1 10  
Db 1 AKEALAL 8

## RESULT 6

US-08-469-660-27  
; Sequence 27, Application US/08469660  
; Patent No. 5876973  
; GENERAL INFORMATION:  
; APPLICANT: Gwynne, David I.; Marchionni, Mark;  
; APPLICANT: McBurney, Robert N.  
; TITLE OF INVENTION: INHIBITORS OF CELL PROLIFERATION,  
; TITLE OF INVENTION: THEIR PREPARATION AND USE  
; NUMBER OF SEQUENCES: 184  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; ZIP: 02111-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
; COMPUTER: IBM  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/469,660  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/011,396  
; FILING DATE: 29-JAN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/984,085  
; FILING DATE: 01-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/951,747  
; FILING DATE: 25-SEP-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/927,337  
; FILING DATE: 10-AUG-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Clark, Paul T.  
; REGISTRATION NUMBER: 30,162  
; REFERENCE/DOCKET NUMBER: 04585/017004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: 200154  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
US-08-469-660-27

Query Match 53.3%; Score 24; DB 2; Length 9;  
Best Local Similarity 75.0%; Pred. No. 2.5e+05;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 AOEALAF1 10  
Db 1 AKEALAL 8

## RESULT 7

US-08-470-335-27  
; Sequence 27, Application US/08470335F  
; Patent No. 6147190  
; GENERAL INFORMATION:  
; APPLICANT: GOODEARL, ANDREW

; APPLICANT: STROOBANT, PAUL  
; APPLICANT: MINGHETTI, LUISA  
; APPLICANT: WATERFIELD, MICHAEL  
; APPLICANT: MARCHIONNI, MARK  
; APPLICANT: CHEN, MARIO S.  
; APPLICANT: HILES, IAN  
; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR  
; FILE REFERENCE: 04585/00200B  
; CURRENT APPLICATION NUMBER: US/08/470,335F  
; CURRENT FILING DATE: 1995-06-06  
; EARLIER APPLICATION NUMBER: 08/036,555  
; EARLIER FILING DATE: 1993-03-24  
; NUMBER OF SEQ ID NOS: 252  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 27  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Bos taurus  
US-08-470-335-27

Query Match 53.3%; Score 24; DB 3; Length 9;  
Best Local Similarity 75.0%; Pred. No. 2.5e+05;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 AOEALAF1 10  
Db 1 AKEALAL 8

## RESULT 8

US-08-735-021-27  
; Sequence 27, Application US/08735021B  
; Patent No. 6194377  
; GENERAL INFORMATION:  
; APPLICANT: GOODEARL, ANDREW  
; APPLICANT: STROOBANT, PAUL  
; APPLICANT: MINGHETTI, LUISA  
; APPLICANT: WATERFIELD, MICHAEL  
; APPLICANT: MARCHIONNI, MARK  
; APPLICANT: CHEN, MARIO S.  
; APPLICANT: HILES, IAN  
; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR  
; FILE REFERENCE: 04585/00200L  
; CURRENT APPLICATION NUMBER: US/08/735,021B  
; CURRENT FILING DATE: 1996-10-22  
; EARLIER APPLICATION NUMBER: 08/472,065  
; EARLIER FILING DATE: 1995-06-06  
; EARLIER APPLICATION NUMBER: 08/036,555  
; EARLIER FILING DATE: 1993-03-24  
; EARLIER APPLICATION NUMBER: 07/965,173  
; EARLIER FILING DATE: 1992-10-23  
; EARLIER APPLICATION NUMBER: 07/940,389  
; EARLIER FILING DATE: 1992-09-03  
; EARLIER APPLICATION NUMBER: 07/907,138  
; EARLIER FILING DATE: 1992-06-30  
; EARLIER APPLICATION NUMBER: 07/863,703  
; EARLIER FILING DATE: 1992-04-03  
; NUMBER OF SEQ ID NOS: 192  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 27  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Bos taurus  
US-08-735-021-27

Query Match 53.3%; Score 24; DB 3; Length 9;  
Best Local Similarity 75.0%; Pred. No. 2.5e+05;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 AOEALAF1 10  
Db 1 AKEALAL 8



Db 1 AKEALAAL 8

RESULT 9  
US-08-734-664A-27  
; Sequence 27, Application US/08734664A  
; Patent No. 6204241  
; GENERAL INFORMATION:  
; APPLICANT: Goodearl, Andrew  
; APPLICANT: Stroobant, Paul  
; APPLICANT: Minghetti, Luisa  
; APPLICANT: Waterfield, Michael  
; APPLICANT: Marchionni, Mark  
; APPLICANT: Chen, Mario  
; APPLICANT: Hiles, Ian  
; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR  
; TITLE OF INVENTION: PREPARATION AND USE  
; NUMBER OF SEQUENCES: 187  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Clark & Elbing LLP  
; STREET: 176 Federal Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IBM Compatible Pentium  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/734,664A  
; FILING DATE: 22-OCT-1996  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/249,322  
; FILING DATE: 26-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/036,555  
; FILING DATE: 24-MAR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/965,173  
; FILING DATE: 23-OCT-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/940,389  
; FILING DATE: 03-SEP-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/907,138  
; FILING DATE: 30-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/863,703  
; FILING DATE: 03-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: UK 91 07566.3  
; FILING DATE: 10-APR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bleker-Brady, Kristina  
; REGISTRATION NUMBER: 39,109  
; REFERENCE/DOCKET NUMBER: 04585/00200J  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 428-0200  
; TELEFAX: (617) 428-7045  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
US-08-734-664A-27

Query Match 53.3%; Score 24; DB 3; Length 9;  
Best Local Similarity 75.0%; Pred. No. 2.5e+05;  
Matches 1; Mismatches 0; Gaps 0;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 AOEALAF 10  
Db 1 AKEALAAL 8

RESULT 10  
US-08-470-339-27  
; Sequence 27, Application US/08470339C  
; Patent No. 6232286  
; GENERAL INFORMATION:  
; APPLICANT: Goodearl, Andrew  
; APPLICANT: Stroobant, Paul  
; APPLICANT: Minghetti, Luisa  
; APPLICANT: Waterfield, Michael  
; APPLICANT: Marchionni, Mark  
; APPLICANT: Chen, Mario S.  
; APPLICANT: Hiles, Ian  
; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR  
; TITLE OF INVENTION: PREPARATION AND USE  
; FILE REFERENCE: 04585/002008  
; CURRENT APPLICATION NUMBER: US/08/470,339C  
; CURRENT FILING DATE: 1995-06-06  
; EARLIER APPLICATION NUMBER: 08/036,555  
; EARLIER FILING DATE: 1993-03-24  
; EARLIER APPLICATION NUMBER: 07/940,389  
; EARLIER FILING DATE: 1992-09-03  
; EARLIER APPLICATION NUMBER: 07/907,138  
; EARLIER FILING DATE: 1992-06-30  
; EARLIER APPLICATION NUMBER: 07/863,703  
; EARLIER FILING DATE: 1992-04-03  
; EARLIER APPLICATION NUMBER: 91 07566.3 GB  
; EARLIER FILING DATE: 1999-04-10  
; NUMBER OF SEQ ID NOS: 226  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 27  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Bos taurus  
US-08-470-339-27

Query Match 53.3%; Score 24; DB 3; Length 9;  
Best Local Similarity 75.0%; Pred. No. 2.5e+05;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 AOEALAF 10  
Db 1 AKEALAAL 8

RESULT 11  
US-08-467-602-27  
; Sequence 27, Application US/08467602C  
; Patent No. 6444642  
; GENERAL INFORMATION:  
; APPLICANT: Sklar, Robert  
; APPLICANT: Marchionni, Mark  
; APPLICANT: Gwynne, David I.  
; TITLE OF INVENTION: METHODS FOR TREATING MUSCLE DISEASES AND  
; TITLE OF INVENTION: DISORDERS  
; FILE REFERENCE: 04585/028003  
; CURRENT APPLICATION NUMBER: US/08/467,602C  
; CURRENT FILING DATE: 1995-06-06  
; EARLIER APPLICATION NUMBER: 08/209,204  
; EARLIER FILING DATE: 1994-03-08  
; EARLIER APPLICATION NUMBER: 08/059,022  
; EARLIER FILING DATE: 1993-05-06  
; NUMBER OF SEQ ID NOS: 420  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 27  
; LENGTH: 9  
; TYPE: PRT

; ORGANISM: Bos taurus  
US-08-467-602-27

Query Match 53.3%; Score 24; DB 4; Length 9;  
Best Local Similarity 75.0%; Pred. No. 2.5e+05;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 AQEALAPL 10  
|:|||||  
Db 1 AKEALAL 8

## RESULT 12

PCT-US94-05083C-27  
; Sequence 27, Application PC/TUS9405083C  
; GENERAL INFORMATION:  
; APPLICANT: Robert Sklar, Mark Marchionni,  
; APPLICANT: David I. Gwynne  
; TITLE OF INVENTION: METHODS FOR ALTERING  
; TITLE OF INVENTION: MUSCLE CONDITION  
; NUMBER OF SEQUENCES: 185  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; ZIP: 02110-2804

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 360  
; MEDIUM TYPE: kb storage  
; COMPUTER: IBM

OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/05083C  
; FILING DATE: 06-MAY-94  
; CLASSIFICATION:

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/209,204  
; FILING DATE: 08-MAR-94  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/059,022  
; FILING DATE: 06-MAY-93  
; ATTORNEY/AGENT INFORMATION:

NAME: Clark, Paul T.  
; REGISTRATION NUMBER: 30,162  
; REFERENCE/DOCKET NUMBER: 04585/028W01  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 200154

INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
PCT-US94-05083C-27

Query Match 53.3%; Score 24; DB 5; Length 9;  
Best Local Similarity 75.0%; Pred. No. 2.5e+05;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 AQEALAPL 10  
|:|||||  
Db 1 AKEALAL 8

## RESULT 13

PCT-US95-06846A-27  
; Sequence 27, Application PC/TUS9506846A  
; GENERAL INFORMATION:  
; APPLICANT: Goodearl, Andrew David; Stroobant, Paul;

; APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;  
; APPLICANT: Chen, Maio Su; Hiles, Ian  
; TITLE OF INVENTION: Glial Mitogenic Factors, Their  
; NUMBER OF SEQUENCES: 178  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Felfe & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
; COMPUTER: IBM  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/06846A  
; FILING DATE: 25-MAY-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/249,322  
; FILING DATE: 26-MAY-1994

CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/036,555  
; FILING DATE: 24-MAR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/965,173  
; FILING DATE: 23-OCT-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/940,389  
; FILING DATE: 03-SEP-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/907,138  
; FILING DATE: 30-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/863,703  
; FILING DATE: 03-APRIL-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: U.K. 91 07566.3  
; FILING DATE: 10-APRIL-1991

ATTORNEY/AGENT INFORMATION:  
; NAME: Hanson, Norman D.  
; REGISTRATION NUMBER: 30,946  
; REFERENCE/DOCKET NUMBER: LUD 5250.5  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 688-9200  
; TELEFAX: (212) 838-3884  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:

LENGTH: 9  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
PCT-US95-06846A-27

Query Match 53.3%; Score 24; DB 5; Length 9;  
Best Local Similarity 75.0%; Pred. No. 2.5e+05;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 AQEALAPL 10  
|:|||||  
Db 1 AKEALAL 8

## RESULT 14

US-08-036-555B-14  
; Sequence 14, Application US/08036555B  
; Patent No. 5530109  
; GENERAL INFORMATION:  
; APPLICANT: Goodearl, Andrew; Stroobant, Paul;  
; APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;

APPLICANT: Chen, Maio Su; Hiles, Ian  
TITLE OF INVENTION: Glial Mitogenic Factors, Their  
PREPARATION AND USE  
NUMBER OF SEQUENCES: 184  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
COMPUTER: IBM  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/036,555B  
FILING DATE: 24-MAR-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/965,173  
FILING DATE: 23-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/940,389  
FILING DATE: 03-SEP-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/907,138  
FILING DATE: 30-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/863,703  
FILING DATE: 10-APRIL-1991  
PRIOR APPLICATION DATA:  
ATTORNEY/AGENT INFORMATION:  
NAME: Tsai, Christine H.  
REGISTRATION NUMBER: 34,266  
REFERENCE/DOCKET NUMBER: LUD 5250.4  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
FEATURE:  
OTHER INFORMATION: Xaa in position 1 is Lysine or Arginine.  
US-08-036-555B-14  
Query Match 53.3%; Score 24; DB 1; Length 10;  
Best Local Similarity 75.0%; Pred. No. 75;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Qy 3 AOEALAF 10  
Db 2 AKEALAL 9  
RESULT 15  
US-08-469-569-14  
Sequence 14, Application US/08469569  
Patent No. 5606032  
GENERAL INFORMATION:  
APPLICANT: Goodearl, Andrew; Stroobant, Paul;  
APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchionni, Mark;  
APPLICANT: Chen, Maio Su; Hiles, Ian  
TITLE OF INVENTION: Glial Mitogenic Factors, Their  
PREPARATION AND USE  
NUMBER OF SEQUENCES: 184  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
COMPUTER: IBM  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/469,569  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/036,555  
FILING DATE: 24-MAR-1993  
APPLICATION NUMBER: 07/965,173  
FILING DATE: 23-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/940,389  
FILING DATE: 03-SEP-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/907,138  
FILING DATE: 30-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/863,703  
FILING DATE: 03-APRIL-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: U.K. 91 07566.3  
FILING DATE: 10-APRIL-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Tsai, Christine H.  
REGISTRATION NUMBER: 34,266  
REFERENCE/DOCKET NUMBER: LUD 5250.4  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
FEATURE:  
OTHER INFORMATION: Xaa in position 1 is Lysine or Arginine.  
US-08-469-569-14  
Query Match 53.3%; Score 24; DB 1; Length 10;  
Best Local Similarity 75.0%; Pred. No. 75;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Qy 3 AOEALAF 10  
Db 2 AKEALAL 9  
Search completed: October 7, 2003, 14:32:03  
Job time : 20.5 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:30:31 ; Search time 41 Seconds  
(without alignments)  
38.589 Million cell updates/sec

Title: US-09-807-512-12  
Perfect score: 45  
Sequence: 1 LMAQEALAF 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues

Total number of hits satisfying chosen parameters: 77217

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	% Match	Query Length	ID	Description
1	27	60.0	9	12	US-10-353-929-157
2	24	53.3	10	9	US-09-834-765-176
3	22	48.9	8	14	US-10-080-797-5
4	22	48.9	9	9	US-09-834-765-452
5	22	48.9	9	9	US-09-834-765-527
6	22	48.9	9	9	US-09-834-765-660
7	22	48.9	9	15	US-10-133-210-114
8	21	46.7	9	9	US-09-834-765-53
9	21	46.7	9	9	US-09-834-765-423
10	21	46.7	9	9	US-09-834-765-647
11	21	46.7	9	11	US-09-865-548A-98
12	21	46.7	9	12	US-10-353-929-28
13	21	46.7	10	9	US-09-834-765-77
14	21	46.7	10	9	US-09-834-765-92
15	21	46.7	10	9	US-09-834-765-411

16	21	46.7	10	10	US-09-995-749A-15
17	21	46.7	10	11	US-09-572-404B-2781
18	21	46.7	10	12	US-09-572-270A-269
19	21	46.7	10	12	US-10-353-929-30
20	21	46.7	10	12	US-10-353-929-31
21	20	44.4	8	9	US-09-884-681-27
22	20	44.4	8	12	US-09-848-107-14
23	20	44.4	8	12	US-10-293-580-27
24	20	44.4	9	9	US-09-810-936-309
25	20	44.4	9	10	US-09-429-755-309
26	20	44.4	9	10	US-09-924-400-309
27	20	44.4	9	15	US-10-212-679-309
28	20	44.4	10	9	US-09-834-765-68
29	20	44.4	10	9	US-09-834-765-279
30	20	44.4	10	11	US-09-572-404B-3178
31	19	42.2	8	10	US-09-982-172-140
32	19	42.2	9	8	US-08-812-393A-47
33	19	42.2	9	9	US-09-834-765-558
34	19	42.2	9	9	US-09-935-682-50
35	19	42.2	9	9	US-09-847-185-18
36	19	42.2	9	10	US-09-909-460-70
37	19	42.2	9	11	US-09-277-074-10
38	19	42.2	9	12	US-10-116-118-40
39	19	42.2	9	12	US-10-150-797-18
40	19	42.2	9	14	US-10-062-257-15
41	19	42.2	9	14	US-10-106-487-18
42	19	42.2	9	15	US-10-001-546-11
43	19	42.2	9	15	US-10-080-013-8
44	19	42.2	9	15	US-10-224-286-18
45	19	42.2	9	15	US-10-091-300-83

ALIGNMENTS

RESULT 1

US-10-353-929-157  
; Sequence 157, Application US/10353929  
; Publication No. US20030175288A1  
; GENERAL INFORMATION:  
; APPLICANT: ITOH, Kyogo  
; TITLE OF INVENTION: Tumor antigen  
; FILE REFERENCE: GP01-1024  
; CURRENT APPLICATION NUMBER: US/10/353,929  
; CURRENT FILING DATE: 2003-01-30  
; PRIOR APPLICATION NUMBER: JP. P2000-231814  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 197  
; SEQ ID NO 157  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Designed peptide based on the amino acid sequence of SEQ ID NO 157  
US-10-353-929-157

Query Match 60.0%; Score 27; DB 12; Length 9;  
Best Local Similarity 75.0%; Pred. No. 5.2e+05;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMAQEALA 8  
Db 1 LLAQEAAA 8

RESULT 2

US-09-834-765-176  
; Sequence 176, Application US/09834765  
; Patent No. US20020055478A1  
; GENERAL INFORMATION:  
; APPLICANT: Mary Paris  
; APPLICANT: Pia M. Challita-Eid

; APPLICANT: Arthur B. Raitano  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT  
; FILE REFERENCE: 129.6USU1  
; CURRENT APPLICATION NUMBER: US/09/834,765  
; CURRENT FILING DATE: 2001-09-21  
; PRIOR APPLICATION NUMBER: 60/197,647  
; PRIOR FILING DATE: 2000-04-12  
; NUMBER OF SEQ ID NOS: 770  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 176  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-834-765-176

Query Match 53.3%; Score 24; DB 9; Length 10;  
Best Local Similarity 60.0%; Pred. No. 96;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LMAQEALAF 10  
| | | : |||  
Db 1 LEAQTNI AFL 10

RESULT 3  
US-10-080-797-5  
; Sequence 5, Application US/10080797  
; Publication No. US20020183253A1  
; GENERAL INFORMATION:  
; APPLICANT: Campochiaro, Peter A.  
; APPLICANT: Dixon, Katharine H.  
; APPLICANT: Brazzell, Romulus K.  
; TITLE OF INVENTION: METHOD FOR TREATING OCULAR  
; TITLE OF INVENTION: NEOVASCULARIZATION  
; FILE REFERENCE: 4-31881A  
; CURRENT APPLICATION NUMBER: US/10/080,797  
; CURRENT FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 21  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Human  
US-10-080-797-5

Query Match 48.9%; Score 22; DB 14; Length 8;  
Best Local Similarity 100.0%; Pred. No. 5.2e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 QEALA 8  
| | | |  
Db 4 QEALA 8

RESULT 4  
US-09-834-765-452  
; Sequence 452, Application US/09834765  
; Patent No. US20020055478A1  
; GENERAL INFORMATION:  
; APPLICANT: Mary Faris  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT  
; FILE REFERENCE: 129.6USU1  
; CURRENT APPLICATION NUMBER: US/09/834,765  
; CURRENT FILING DATE: 2001-09-21  
; PRIOR APPLICATION NUMBER: 60/197,647  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 176  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-834-765-176

; CURRENT FILING DATE: 2001-09-21  
; PRIOR APPLICATION NUMBER: 60/197,647  
; PRIOR FILING DATE: 2000-04-12  
; NUMBER OF SEQ ID NOS: 770  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 452  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-834-765-452

Query Match 48.9%; Score 22; DB 9; Length 9;  
Best Local Similarity 62.5%; Pred. No. 5.2e+05;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 AQEALAF 10  
| | | : |||  
Db 2 AQTNI AFL 9

RESULT 5  
US-09-834-765-527  
; Sequence 527, Application US/09834765  
; Patent No. US20020055478A1  
; GENERAL INFORMATION:  
; APPLICANT: Mary Faris  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT  
; FILE REFERENCE: 129.6USU1  
; CURRENT APPLICATION NUMBER: US/09/834,765  
; CURRENT FILING DATE: 2001-09-21  
; PRIOR APPLICATION NUMBER: 60/197,647  
; PRIOR FILING DATE: 2000-04-12  
; NUMBER OF SEQ ID NOS: 770  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 527  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-834-765-527

Query Match 48.9%; Score 22; DB 9; Length 9;  
Best Local Similarity 62.5%; Pred. No. 5.2e+05;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 AQEALAF 10  
| | | : |||  
Db 2 AQTNI AFL 9

RESULT 6  
US-09-834-765-660  
; Sequence 660, Application US/09834765  
; Patent No. US20020055478A1  
; GENERAL INFORMATION:  
; APPLICANT: Mary Faris  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT  
; FILE REFERENCE: 129.6USU1  
; CURRENT APPLICATION NUMBER: US/09/834,765  
; CURRENT FILING DATE: 2001-09-21  
; PRIOR APPLICATION NUMBER: 60/197,647  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 527  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-834-765-527

; NUMBER OF SEQ ID NOS: 770  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 660  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-834-765-660

Query Match 48.9%; Score 22; DB 9; Length 9;  
Best Local Similarity 62.5%; Pred. No. 5.2e+05;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 AOEALAF 10  
|| :|||  
Db 2 AOTNIAFL 9

## RESULT 7

US-10-133-210-114  
; Sequence 114, Application US/10133210  
; Publication No. US20030103964A1  
; GENERAL INFORMATION:  
; APPLICANT: DeLisi, Charles  
; APPLICANT: Berzofsky, Jay  
; APPLICANT: Gulukota, Kamalakara  
; APPLICANT: Vaccaro, Dennis  
; APPLICANT: Weng, Zhiping  
; APPLICANT: Zhang, Chao  
; TITLE OF INVENTION: METHODS FOR DESIGNING MOLECULAR CONJUGATES AND  
; TITLE OF INVENTION: COMPOSITIONS THEREOF  
; FILE REFERENCE: BU-035AX  
; CURRENT APPLICATION NUMBER: US/10/133,210  
; CURRENT FILING DATE: 2002-04-26  
; NUMBER OF SEQ ID NOS: 281  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 114  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-10-133-210-114

Query Match 48.9%; Score 22; DB 15; Length 9;  
Best Local Similarity 71.4%; Pred. No. 5.2e+05;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 QEALAF 10  
:|:||||  
Db 3 REDLAF 9

## RESULT 8

US-09-834-765-53  
; Sequence 53, Application US/09834765  
; Patent No. US20020055478A1  
; GENERAL INFORMATION:  
; APPLICANT: Mary Faris  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT  
; TITLE OF INVENTION: AND DETECTION OF CANCER  
; FILE REFERENCE: 129.6USU1  
; CURRENT APPLICATION NUMBER: US/09/834,765  
; CURRENT FILING DATE: 2001-09-21  
; PRIOR FILING DATE: 2000-04-12  
; NUMBER OF SEQ ID NOS: 770  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 53

; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-834-765-53

Query Match 46.7%; Score 21; DB 9; Length 9;  
Best Local Similarity 62.5%; Pred. No. 5.2e+05;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 AOEALAF 10  
|:||||:  
Db 2 ABEALSDL 9

## RESULT 9

US-09-834-765-423  
; Sequence 423, Application US/09834765  
; Patent No. US20020055478A1  
; GENERAL INFORMATION:  
; APPLICANT: Mary Faris  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT  
; TITLE OF INVENTION: AND DETECTION OF CANCER  
; FILE REFERENCE: 129.6USU1  
; CURRENT APPLICATION NUMBER: US/09/834,765  
; CURRENT FILING DATE: 2001-09-21  
; PRIOR FILING DATE: 2000-04-12  
; NUMBER OF SEQ ID NOS: 770  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 423  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-834-765-423

Query Match 46.7%; Score 21; DB 9; Length 9;  
Best Local Similarity 62.5%; Pred. No. 5.2e+05;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 AOEALAF 10  
|:||||:  
Db 2 ABEALSDL 9

## RESULT 10

US-09-834-765-647  
; Sequence 647, Application US/09834765  
; Patent No. US20020055478A1  
; GENERAL INFORMATION:  
; APPLICANT: Mary Faris  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT  
; TITLE OF INVENTION: AND DETECTION OF CANCER  
; FILE REFERENCE: 129.6USU1  
; CURRENT APPLICATION NUMBER: US/09/834,765  
; CURRENT FILING DATE: 2001-09-21  
; PRIOR FILING DATE: 2000-04-12  
; NUMBER OF SEQ ID NOS: 770  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 647  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens

US-09-834-765-647

Query Match 46.7%; Score 21; DB 9; Length 9;  
Best Local Similarity 62.5%; Pred. No. 5.2e+05;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 AQEALAFI 10  
|:|:|:|  
Db 2 AEEALSDF 9

RESULT 11

US-09-865-548A-98

; Sequence 98, Application US/09865548A

; Publication No. US20030096298A1

; GENERAL INFORMATION:

; APPLICANT: Barnea, Eilon

; APPLICANT: Beer, Ilan

; APPLICANT: Ziv, Tamar

; APPLICANT: Admon, Arie

; TITLE OF INVENTION: METHOD OF IDENTIFYING PEPTIDES CAPABLE OF BINDING TO MHC MOLECULES

; FILE REFERENCE: 01/22080

; CURRENT APPLICATION NUMBER: US/09/865,548A

; CURRENT FILING DATE: 2001-05-16

; PRIOR APPLICATION NUMBER: US 60/290,958

; PRIOR FILING DATE: 2001-05-16

; NUMBER OF SEQ ID NOS: 204

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 98

; LENGTH: 9

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: synthetic peptide

US-09-865-548A-98

Query Match

46.7%; Score 21; DB 11; Length 9;

Best Local Similarity 55.6%; Pred. No. 5.2e+05;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 MAQALAFI 10  
|:|:|:|  
Db 1 MAPQALLL 9

RESULT 12

US-10-353-929-28

; Sequence 28, Application US/10353929

; Publication No. US20030175288A1

; GENERAL INFORMATION:

; APPLICANT: ITOH, Kyogo

; TITLE OF INVENTION: Tumor antigen

; FILE REFERENCE: GP01-1024

; CURRENT APPLICATION NUMBER: US/10/353,929

; CURRENT FILING DATE: 2003-01-30

; PRIOR APPLICATION NUMBER: JP P2000-231814

; PRIOR FILING DATE: 2000-07-31

; NUMBER OF SEQ ID NOS: 197

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 28

; LENGTH: 9

; TYPE: PRT

; ORGANISM: Artificial

; FEATURE:

; OTHER INFORMATION: Designed peptide having an ability to activate HLA-A2 restricted

; OTHER INFORMATION: Cytotoxic T lymphocytes

US-10-353-929-28

Query Match

46.7%; Score 21; DB 12; Length 9;

Best Local Similarity 71.4%; Pred. No. 5.2e+05;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LMAQEAL 7  
|:|:|:|  
Db 2 LMLQNAL 8

RESULT 13

US-09-834-765-77

; Sequence 77, Application US/09834765

; Patent No. US20020055478A1

; GENERAL INFORMATION:

; APPLICANT: Mary Paris

; APPLICANT: Pia M. Challita-Eid

; APPLICANT: Arthur B. Raitano

; APPLICANT: Steve Chappell Mitchell

; APPLICANT: Daniel E.H. Afar

; APPLICANT: Aya Jakobovits

; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT

; FILE REFERENCE: 129.6USU1

; CURRENT APPLICATION NUMBER: US/09/834,765

; CURRENT FILING DATE: 2001-09-21

; PRIOR APPLICATION NUMBER: 60/197,647

; PRIOR FILING DATE: 2000-04-12

; NUMBER OF SEQ ID NOS: 770

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 77

; LENGTH: 10

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-834-765-77

Query Match

46.7%; Score 21; DB 9; Length 10;

Best Local Similarity 62.5%; Pred. No. 4e+02;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 AQEALAFI 10  
|:|:|:|  
Db 2 AEEALSDF 9

RESULT 14

US-09-834-765-92

; Sequence 92, Application US/09834765

; Patent No. US20020055478A1

; GENERAL INFORMATION:

; APPLICANT: Mary Paris

; APPLICANT: Pia M. Challita-Eid

; APPLICANT: Arthur B. Raitano

; APPLICANT: Steve Chappell Mitchell

; APPLICANT: Daniel E.H. Afar

; APPLICANT: Aya Jakobovits

; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT

; FILE REFERENCE: 129.6USU1

; CURRENT APPLICATION NUMBER: US/09/834,765

; CURRENT FILING DATE: 2001-09-21

; PRIOR APPLICATION NUMBER: 60/197,647

; PRIOR FILING DATE: 2000-04-12

; NUMBER OF SEQ ID NOS: 770

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 92

; LENGTH: 10

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-834-765-92

Query Match

46.7%; Score 21; DB 9; Length 10;

Best Local Similarity 62.5%; Pred. No. 4e+02;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 AQEALAFI 10  
|:|:|:|  
Db 1 AEEALSDF 8



```
RESULT 15
US-09-834-765-411
; Sequence 411, Application US/09834765
; Patent No. US20020055478A1
; GENERAL INFORMATION:
; APPLICANT: Mary Faris
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Arthur B. Raitano
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT
; FILE REFERENCE: 129.6USU1
; CURRENT APPLICATION NUMBER: US/09/834,765
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: 60/197,647
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 770
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 411
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-765-411

Query Match      46.7%; Score 21; DB 9; Length 10;
Best Local Similarity 62.5%; Pred. No. 4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      3 AOEALAFI 10
Db      2 AEEALSDF 9

Search completed: October 7, 2003, 14:43:29
Job time : 41 secs
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:27:56 ; Search time 276.5 Seconds  
(without alignments)  
32.908 Million cell updates/sec

Title: US-09-807-512-12  
Perfect score: 45  
Sequence: 1 LMAQELALFL 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues  
Total number of hits satisfying chosen parameters: 566550

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main:\*

1: /cgn2\_6/ptodata/1/paa/PCTRUS\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US080\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep.\*  
8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep.\*  
9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep.\*  
10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep.\*  
11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep.\*  
12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep.\*  
13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep.\*  
14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep.\*  
15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep.\*  
16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep.\*  
17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep.\*  
18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep.\*  
19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep.\*  
20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep.\*  
21: /cgn2\_6/ptodata/1/paa/US097A\_COMB.pep.\*  
22: /cgn2\_6/ptodata/1/paa/US097B\_COMB.pep.\*  
23: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep.\*  
24: /cgn2\_6/ptodata/1/paa/US099A\_COMB.pep.\*  
25: /cgn2\_6/ptodata/1/paa/US099B\_COMB.pep.\*  
26: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep.\*  
27: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep.\*  
28: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep.\*  
29: /cgn2\_6/ptodata/1/paa/US103\_COMB.pep.\*  
30: /cgn2\_6/ptodata/1/paa/US104\_COMB.pep.\*  
31: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep.\*  
32: /cgn2\_6/ptodata/1/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	100.0	10	23 US-09-807-512-12	Sequence 12, Appl

2	27	60.0	9	29	US-10-353-929-157	Sequence 157, Appl
3	24	53.3	9	4	US-08-011-396A-27	Sequence 27, Appl
4	24	53.3	9	6	US-08-209-204A-27	Sequence 27, Appl
5	24	53.3	9	6	US-08-209-204B-27	Sequence 27, Appl
6	24	53.3	9	6	US-08-209-204D-27	Sequence 27, Appl
7	24	53.3	9	8	US-08-461-097-27	Sequence 27, Appl
8	24	53.3	9	8	US-08-461-097A-27	Sequence 27, Appl
10	24	53.3	9	8	US-08-461-097-27	Sequence 27, Appl
11	24	53.3	9	8	US-08-466-446-27	Sequence 27, Appl
12	24	53.3	9	8	US-08-467-459-27	Sequence 27, Appl
13	24	53.3	9	8	US-08-467-459A-27	Sequence 27, Appl
14	24	53.3	9	8	US-08-467-470-27	Sequence 27, Appl
15	24	53.3	9	8	US-08-467-906-27	Sequence 27, Appl
16	24	53.3	9	8	US-08-468-731-27	Sequence 27, Appl
17	24	53.3	9	8	US-08-468-731A-27	Sequence 27, Appl
18	24	53.3	9	8	US-08-468-731-27	Sequence 27, Appl
19	24	53.3	9	8	US-08-468-526-27	Sequence 27, Appl
20	24	53.3	9	8	US-08-469-549-27	Sequence 27, Appl
21	24	53.3	9	8	US-08-469-549-27	Sequence 27, Appl
22	24	53.3	9	8	US-08-470-335-27	Sequence 27, Appl
23	24	53.3	9	8	US-08-470-335B-27	Sequence 27, Appl
24	24	53.3	9	8	US-08-470-335-27	Sequence 27, Appl
25	24	53.3	9	8	US-08-471-148-27	Sequence 27, Appl
26	24	53.3	9	8	US-08-471-833-27	Sequence 27, Appl
27	24	53.3	9	8	US-08-471-833-27	Sequence 27, Appl
28	24	53.3	9	8	US-08-471-855-27	Sequence 27, Appl
29	24	53.3	9	8	US-08-472-008-27	Sequence 27, Appl
30	24	53.3	9	8	US-08-472-065-27	Sequence 27, Appl
31	24	53.3	9	8	US-08-472-065F-27	Sequence 27, Appl
32	24	53.3	9	9	US-08-535-200-27	Sequence 27, Appl
33	24	53.3	9	9	US-08-535-200A-27	Sequence 27, Appl
34	24	53.3	9	11	US-08-734-592-27	Sequence 27, Appl
35	24	53.3	9	11	US-08-734-665-27	Sequence 27, Appl
36	24	53.3	9	11	US-08-734-665A-27	Sequence 27, Appl
37	24	53.3	9	11	US-08-734-665-27	Sequence 27, Appl
38	24	53.3	9	11	US-08-734-666-27	Sequence 27, Appl
39	24	53.3	9	11	US-08-735-010-27	Sequence 27, Appl
40	24	53.3	9	11	US-08-736-019-27	Sequence 27, Appl
41	24	53.3	9	11	US-08-736-019-27	Sequence 27, Appl
42	24	53.3	9	11	US-08-736-070A-27	Sequence 27, Appl
43	24	53.3	9	11	US-08-736-070-27	Sequence 27, Appl
44	24	53.3	9	11	US-08-736-070-27	Sequence 27, Appl
45	24	53.3	10	4	US-08-011-396A-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1  
US-09-807-512-12  
; Sequence 12, Application US/09807512  
; GENERAL INFORMATION:  
; APPLICANT: Schrier, Peter I.  
; APPLICANT: Aarnoudse, Corlien  
; APPLICANT: Heider, Karl-Heinz  
; APPLICANT: Klade, Christoph  
; TITLE OF INVENTION: Camel, An Alternative Translation Product of the Tumor  
; TITLE OF INVENTION: Antigen-Lage 1  
; FILE REFERENCE: 0652.2200000  
; CURRENT APPLICATION NUMBER: US/09/807,512  
; CURRENT FILING DATE: 2001-04-16  
; PRIOR APPLICATION NUMBER: PCT/EP99/07832  
; PRIOR FILING DATE: 1999-10-15  
; PRIOR APPLICATION NUMBER: EP 98119583.7  
; PRIOR FILING DATE: 1998-10-16  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 12  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-807-512-12

Query Match 100.0%; Score 45; DB 23; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.039; 0; Indels 0; Gaps 0;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LMAQEALAF 10  
Db 1 LMAQEALAF 10

## RESULT 2

US-10-353-929-157  
Sequence 157, Application US/10353929

GENERAL INFORMATION:

APPLICANT: ITOH, Kyoto

TITLE OF INVENTION: Tumor antigen

FILE REFERENCE: GP01-1024

CURRENT APPLICATION NUMBER: US/10/353,929

CURRENT FILING DATE: 2003-01-30

PRIOR APPLICATION NUMBER: JP P2000-231814

PRIOR FILING DATE: 2000-07-31

NUMBER OF SEQ ID NOS: 197

SOFTWARE: Patent version 3.1

SEQ ID NO 157

LENGTH: 9

TYPE: PRT

ORGANISM: Artificial

FEATURE:

OTHER INFORMATION: Designed peptide based on the amino acid sequence of SEQ ID NO:50  
US-10-353-929-157

Query Match 60.0%; Score 27; DB 29; Length 9;  
Best Local Similarity 75.0%; Pred. No. 5.3e+06; 1; Indels 0; Gaps 0;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LMAQEAL 8  
Db 1 LLAQEAA 8

## RESULT 3

US-08-011-396A-27

Sequence 27, Application US/08011396A

GENERAL INFORMATION:

APPLICANT: Gwynne, David I.; Marchionni, Mark;

APPLICANT: McBurney, Robert N.

TITLE OF INVENTION: INHIBITORS OF CELL PROLIFERATION,

TITLE OF INVENTION: THEIR PREPARATION AND USE

NUMBER OF SEQUENCES: 184

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson

STREET: 225 Franklin Street

CITY: Boston

STATE: Massachusetts

ZIP: 02211-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/011.396A

FILING DATE: 29-JAN-1993

CLASSIFICATION: 436

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/984,085

FILING DATE: 01-DEC-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/951,747

FILING DATE: 25-SEP-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/927,337

FILING DATE: 10-AUG-1992

ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.

REGISTRATION NUMBER: 30,162

REFERENCE/DOCKET NUMBER: 04585/017004

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 542-5070

TELEFAX: 200154

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:

LENGTH: 9

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

US-08-011-396A-27

Query Match 53.3%; Score 24; DB 4; Length 9;  
Best Local Similarity 75.0%; Pred. No. 5.3e+06; 1; Indels 0; Gaps 0;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 AQEALAF 10  
Db 1 AQEALAF 8

## RESULT 4

US-08-209-204A-27

Sequence 27, Application US/08209204A

GENERAL INFORMATION:

APPLICANT: Robert Sklar, Mark Marchionni,

APPLICANT: David I. Gwynne

TITLE OF INVENTION: METHODS FOR TREATING

NUMBER OF SEQUENCES: 184

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson

STREET: 225 Franklin Street

CITY: Boston

STATE: Massachusetts

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/209,204A

FILING DATE: 08-MAR-1994

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/059,022

FILING DATE: 06-MAY-93

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Clark, Paul T.

REGISTRATION NUMBER: 30,162

REFERENCE/DOCKET NUMBER: 04585/028001

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 542-5070

TELEFAX: 200154

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:

LENGTH: 9

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

US-08-209-204A-27

Query Match 53.3%; Score 24; DB 6; Length 9;  
Best Local Similarity 75.0%; Pred. No. 5.3e+06; 1; Indels 0; Gaps 0;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 AQEALAF 10



APPLICANT: Sklar, Robert  
APPLICANT: Marchionni, Mark  
APPLICANT: Gwynne, David  
TITLE OF INVENTION: METHODS FOR TREATING MUSCLE  
DISEASES AND DISORDERS  
NUMBER OF SEQUENCES: 383  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/461.097A  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/059,022  
FILING DATE: 06-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/209,204  
FILING DATE: 08-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Bleker-Brady, Kristina  
REGISTRATION NUMBER: 39,109  
REFERENCE/DOCKET NUMBER: 04585/028002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX:  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-461-097A-27

Query Match 53.3%; Score 24; DB 8; Length 9;  
Best Local Similarity 75.0%; Pred. No. 5.3e+06;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 AOEALAF 10  
DB 1 AKEALAL 8

RESULT 9  
US-08-461-097-27  
Sequence 27, Application US/08461097C  
GENERAL INFORMATION:  
APPLICANT: Sklar, Robert  
APPLICANT: Marchionni, Mark  
APPLICANT: Gwynne, David I.  
TITLE OF INVENTION: METHODS FOR TREATING MUSCLE DISEASES AND  
DISORDERS  
FILE REFERENCE: 04585/028002  
CURRENT APPLICATION NUMBER: US/08/461.097C  
CURRENT FILING DATE: 1995-06-05  
EARLIER APPLICATION NUMBER: 08/209,204  
EARLIER FILING DATE: 1994-03-08  
EARLIER APPLICATION NUMBER: 08/059,022  
EARLIER FILING DATE: 1993-05-06  
NUMBER OF SEQ ID NOS: 420  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 27  
LENGTH: 9

TYPE: PRT  
ORGANISM: Bos taurus  
US-08-461-097-27

Query Match 53.3%; Score 24; DB 8; Length 9;  
Best Local Similarity 75.0%; Pred. No. 5.3e+06;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 AOEALAF 10  
DB 1 AKEALAL 8

RESULT 10  
US-08-466-446-27  
Sequence 27, Application US/08466446  
GENERAL INFORMATION:  
APPLICANT: Gwynne, David I.; Marchionni, Mark;  
APPLICANT: McBurney, Robert N.  
TITLE OF INVENTION: INHIBITORS OF CELL PROLIFERATION,  
THEIR PREPARATION AND USE  
NUMBER OF SEQUENCES: 184  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
ZIP: 02111-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
COMPUTER: IBM  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/466,446  
FILING DATE:  
CLASSIFICATION: 436  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/011,396  
FILING DATE: 29-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/984,085  
FILING DATE: 01-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/951,747  
FILING DATE: 25-SEP-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/927,337  
FILING DATE: 10-AUG-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 04585/017004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: 200154  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-466-446-27

Query Match 53.3%; Score 24; DB 8; Length 9;  
Best Local Similarity 75.0%; Pred. No. 5.3e+06;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 AOEALAF 10  
DB 1 AKEALAL 8

```
RESULT 11
US-08-467-459-27
; Sequence 27, Application US/08467459
; GENERAL INFORMATION:
; APPLICANT: Gwynne, David I.; Marchionni, Mark;
; APPLICANT: McBurney, Robert N.
; TITLE OF INVENTION: INHIBITORS OF CELL PROLIFERATION,
; TITLE OF INVENTION: THEIR PREPARATION AND USE
; NUMBER OF SEQUENCES: 184
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; ZIP: 02111-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,459
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/011,396
; FILING DATE: 29-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/984,085
; FILING DATE: 01-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/951,747
; FILING DATE: 25-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/927,337
; FILING DATE: 10-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 04585/017004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: 200154
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9
; TYPE: amino acid
; STRANDEDNESS: linear
; TOPOLOGY: linear
US-08-467-459-27

Query Match 53.3%; Score 24; DB 8; Length 9;
Best Local Similarity 75.0%; Pred. No. 5.3e+06;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 AQEALAF 10
Db 1 AKEALAA 8

RESULT 12
US-08-467-459A-27
; Sequence 27, Application US/08467459A
; GENERAL INFORMATION:
; APPLICANT: Gwynne, David I
; APPLICANT: MARCHIONNI, MARK
; APPLICANT: MCBURNEY, ROBERT N
; TITLE OF INVENTION: INHIBITORS OF CELL PROLIFERATION,
; TITLE OF INVENTION: THEIR PREPARATION AND USE
; NUMBER OF SEQUENCES: 188
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
```

```
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,459A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/011,396
; FILING DATE: 29-JAN-1993
; APPLICATION NUMBER: 07/984,085
; FILING DATE: 01-DEC-1992
; APPLICATION NUMBER: 07/951,747
; FILING DATE: 25-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Bieker-Brady, Kristina
; REGISTRATION NUMBER: 39,109
; REFERENCE/DOCKET NUMBER: 04585/017008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-467-459A-27

Query Match 53.3%; Score 24; DB 8; Length 9;
Best Local Similarity 75.0%; Pred. No. 5.3e+06;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 AQEALAF 10
Db 1 AKEALAA 8

RESULT 13
US-08-467-459-27
; Sequence 27, Application US/08467459C
; GENERAL INFORMATION:
; APPLICANT: Gwynne, David I
; APPLICANT: MARCHIONNI, MARK
; APPLICANT: MCBURNEY, ROBERT N.
; TITLE OF INVENTION: METHOD FOR MAKING ANTIBODIES WHICH
; TITLE OF INVENTION: SPECIFICALLY BIND TO GLI3 GROWTH FACTORS (AS AMENDED)
; FILE REFERENCE: 04585/017008
; CURRENT APPLICATION NUMBER: US/08/467,459C
; CURRENT FILING DATE: 1995-06-06
; EARLIER APPLICATION NUMBER: 08/011,396
; EARLIER FILING DATE: 1993-01-29
; EARLIER APPLICATION NUMBER: 07/984,085
; EARLIER FILING DATE: 1992-12-01
; EARLIER APPLICATION NUMBER: 07/951,747
; EARLIER FILING DATE: 1992-09-25
; NUMBER OF SEQ ID NOS: 640
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Bos taurus
US-08-467-459-27

Query Match 53.3%; Score 24; DB 8; Length 9;
```

Best Local Similarity 75.0%; Pred. No. 5.3e+06;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 AQEALAF 10  
1:|||||  
Db 1 AKEALAL 8

## RESULT 14

US-08-467-470-27  
; Sequence 27, Application US/08467470  
; GENERAL INFORMATION:  
; APPLICANT: Robert Sklar, Mark Marchionni,  
; APPLICANT: David I. Gwynne  
; TITLE OF INVENTION: METHODS FOR TREATING  
; TITLE OF INVENTION: MUSCLE DISEASES AND DISORDERS  
; NUMBER OF SEQUENCES: 184  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
; COMPUTER: IBM  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/467,470  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/209,204  
; FILING DATE: 08-MAR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/059,022  
; FILING DATE: 06-MAY-93  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Clark, Paul T.  
; REGISTRATION NUMBER: 30,162  
; REFERENCE/DOCKET NUMBER: 04585/028005  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: 200154  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
US-08-467-470-27

Query Match 53.3%; Score 24; DB 8; Length 9;  
Best Local Similarity 75.0%; Pred. No. 5.3e+06;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 AQEALAF 10  
1:|||||  
Db 1 AKEALAL 8

## RESULT 15

US-08-467-906-27  
; Sequence 27, Application US/08467906  
; GENERAL INFORMATION:  
; APPLICANT: Gwynne, David I.; Marchionni, Mark;  
; APPLICANT: McBurney, Robert N.  
; TITLE OF INVENTION: INHIBITORS OF CELL PROLIFERATION,  
; TITLE OF INVENTION: THEIR PREPARATION AND USE  
; NUMBER OF SEQUENCES: 184  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson

STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
ZIP: 0211-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
COMPUTER: IBM  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/467,906  
FILING DATE:  
CLASSIFICATION: 436  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/011,396  
FILING DATE: 29-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/984,085  
FILING DATE: 01-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/951,747  
FILING DATE: 25-SEP-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/927,337  
FILING DATE: 10-AUG-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 04585/017004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: 200154  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-467-906-27

Query Match 53.3%; Score 24; DB 8; Length 9;  
Best Local Similarity 75.0%; Pred. No. 5.3e+06;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 AQEALAF 10  
1:|||||  
Db 1 AKEALAL 8

Search completed: October 7, 2003, 14:41:24  
Job time : 277.5 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:28:26 ; Search time 14.5 seconds  
(without alignments)  
20.087 Million cell updates/sec

Title: US-09-807-512-12

Perfect score: 45

Sequence: 1 LMAQBALAFL 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 136899 seqs, 29126274 residues

Total number of hits satisfying chosen parameters: 23559

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*

- 1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query %	Score	Match	Length	ID	Description
1	19	42.2	8	1	PCT-US03-27706-385	Sequence 385, Appl
2	19	42.2	8	5	US-09-883-859-46	Sequence 46, Appl
3	19	42.2	8	6	US-10-657-022-385	Sequence 385, Appl
4	19	42.2	9	1	PCT-US03-27706-384	Sequence 384, Appl
5	19	42.2	9	5	US-09-872-186-6	Sequence 6, Appl
6	19	42.2	9	5	US-09-856-988B-29	Sequence 29, Appl
7	19	42.2	9	6	US-10-647-005-27	Sequence 27, Appl
8	19	42.2	9	6	US-10-657-022-384	Sequence 384, Appl
9	19	42.2	9	7	US-60-485-1717-87	Sequence 87, Appl
10	19	42.2	10	1	PCT-US03-27706-383	Sequence 383, Appl
11	19	42.2	10	1	PCT-US03-27706-386	Sequence 386, Appl
12	19	42.2	10	1	PCT-US03-27706-387	Sequence 387, Appl
13	19	42.2	10	5	US-09-239-043D-19	Sequence 19, Appl
14	19	42.2	10	5	US-09-239-043D-696	Sequence 696, Appl
15	19	42.2	10	5	US-09-239-043D-1535	Sequence 1535, Appl
16	19	42.2	10	5	US-09-239-043D-1798	Sequence 1798, Appl
17	19	42.2	10	6	US-10-657-022-383	Sequence 383, Appl
18	19	42.2	10	6	US-10-657-022-386	Sequence 386, Appl
19	19	42.2	10	6	US-10-657-022-387	Sequence 387, Appl
20	19	42.2	10	6	US-10-658-180-213	Sequence 213, Appl
21	19	42.2	10	6	US-10-658-180-217	Sequence 217, Appl
22	19	42.2	10	6	US-10-363-990-19	Sequence 19, Appl
23	19	42.2	10	6	US-10-363-990-731	Sequence 731, Appl
24	19	42.2	10	6	US-10-363-990-1824	Sequence 1824, Appl
25	19	42.2	10	6	US-10-363-990-2125	Sequence 2125, Appl
26	18	40.0	8	5	US-09-390-061D-250	Sequence 250, Appl

27	18	40.0	8	5	US-09-390-061D-2049	Sequence 2049, Appl
28	18	40.0	9	1	PCT-US03-23503-29	Sequence 29, Appl
29	18	40.0	9	1	PCT-US03-27706-66	Sequence 66, Appl
30	18	40.0	9	5	US-09-390-061D-613	Sequence 613, Appl
31	18	40.0	9	5	US-09-390-061D-2050	Sequence 2050, Appl
32	18	40.0	9	5	US-09-390-061D-2238	Sequence 2238, Appl
33	18	40.0	9	5	US-09-390-061D-2254	Sequence 2254, Appl
34	18	40.0	9	5	US-09-390-061D-2697	Sequence 2697, Appl
35	18	40.0	9	5	US-09-390-061D-3417	Sequence 3417, Appl
36	18	40.0	9	5	US-09-390-061D-3418	Sequence 3418, Appl
37	18	40.0	9	5	US-09-453-174-14	Sequence 14, Appl
38	18	40.0	9	6	US-10-657-022-66	Sequence 66, Appl
39	18	40.0	10	1	PCT-US03-27706-67	Sequence 67, Appl
40	18	40.0	10	5	US-09-390-061D-2221	Sequence 2221, Appl
41	18	40.0	10	5	US-09-390-061D-2239	Sequence 2239, Appl
42	18	40.0	10	5	US-09-390-061D-2255	Sequence 2255, Appl
43	18	40.0	10	5	US-09-390-061D-2882	Sequence 2882, Appl
44	18	40.0	10	5	US-09-390-061D-2897	Sequence 2897, Appl
45	18	40.0	10	5	US-09-694-497B-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1  
PCT-US03-27706-385  
; Sequence 385, Application PC/TUS0327706  
; GENERAL INFORMATION:  
; APPLICANT: MANNKIND CORPORATION  
; APPLICANT: SIMARD, John J. L.  
; APPLICANT: DIAMOND, David C.  
; APPLICANT: LIU, Liping  
; APPLICANT: LIU, Zheng  
; TITLE OF INVENTION: EPITOPE SEQUENCES  
; FILE REFERENCE: MANNK.032VPC  
; CURRENT APPLICATION NUMBER: PCT/US03/27706  
; CURRENT FILING DATE: 2003-09-05  
; PRIOR APPLICATION NUMBER: US 60/409123  
; PRIOR FILING DATE: 2002-09-06  
; NUMBER OF SEQ ID NOS: 610  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 385  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Homosapiens  
PCT-US03-27706-385

Query Match 42.2%; Score 19; DB 1; Length 8;  
Best Local Similarity 66.7%; Pred. No. 1.2e+05;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 5 EALAF 10  
| | | |  
DB 3 ERLAY 8

RESULT 2  
US-09-883-859-46  
; Sequence 46, Application US/09883859  
; GENERAL INFORMATION:  
; APPLICANT: Sicheiri et al.  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURE OF A STERILE ALPHA MOTIF DOMAIN  
; CURRENT APPLICATION NUMBER: US/09/883,859  
; CURRENT FILING DATE: 2001-06-18  
; PRIOR APPLICATION NUMBER: PCT/CA99/01209  
; PRIOR FILING DATE: 1999-12-17  
; NUMBER OF SEQ ID NOS: 126  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 46  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence

; FEATURE:  
; OTHER INFORMATION: peptide molecule that modulates SAM domain function  
US-09-883-859-46

Query Match 42.2%; Score 19; DB 5; Length 8;  
Best Local Similarity 57.1%; Pred. No. 1.2e+05;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 MAQALA 8  
| : | : |  
DB 2 MSQDDLA 8

RESULT 3  
US-10-657-022-385  
; Sequence 385, Application US/10657022  
; GENERAL INFORMATION:  
; APPLICANT: Simard, John J. L.  
; APPLICANT: Diamond, David C.  
; APPLICANT: Liu, Liping  
; APPLICANT: Liu, Zheng  
; TITLE OF INVENTION: EPITOPE SEQUENCES  
; FILE REFERENCE: MANK.032A  
; CURRENT APPLICATION NUMBER: US/10/657.022  
; CURRENT FILING DATE: 2003-09-04  
; PRIOR APPLICATION NUMBER: 60/409123  
; PRIOR FILING DATE: 2002-09-06  
; NUMBER OF SEQ ID NOS: 610  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 385  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Homosapiens  
US-10-657-022-385

Query Match 42.2%; Score 19; DB 6; Length 8;  
Best Local Similarity 66.7%; Pred. No. 1.2e+05;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 EALAF 10  
| : | : |  
DB 3 ERLAYL 8

RESULT 4  
PCT-US03-27706-384  
; Sequence 384, Application PC/TUS0327706  
; GENERAL INFORMATION:  
; APPLICANT: MANKIND CORPORATION  
; APPLICANT: SIMARD, John J. L.  
; APPLICANT: DIAMOND, David C.  
; APPLICANT: LIU, Liping  
; APPLICANT: LIU, Zheng  
; TITLE OF INVENTION: EPITOPE SEQUENCES  
; FILE REFERENCE: MANK.032VPC  
; CURRENT APPLICATION NUMBER: PCT/US03/27706  
; CURRENT FILING DATE: 2003-09-05  
; PRIOR APPLICATION NUMBER: US 60/409123  
; PRIOR FILING DATE: 2002-09-06  
; NUMBER OF SEQ ID NOS: 610  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 384  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homosapiens  
PCT-US03-27706-384

Query Match 42.2%; Score 19; DB 1; Length 9;  
Best Local Similarity 66.7%; Pred. No. 1.2e+05;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 EALAF 10  
| : | : |

Db 4 ERLAYL 9

## RESULT 5

US-09-872-186-6  
; Sequence 6, Application US/09872186  
; GENERAL INFORMATION:  
; APPLICANT: John R. Subjeck  
; APPLICANT: Robert A. Henderson  
; APPLICANT: Elizabeth A. Repasky  
; APPLICANT: Latif Kazim  
; APPLICANT: Xiang-Yang Wang  
; TITLE OF INVENTION: STRESS PROTEIN COMPOSITIONS AND METHODS  
; TITLE OF INVENTION: FOR PREVENTION AND TREATMENT OF CANCER AND INFECTIOUS  
; TITLE OF INVENTION: DISEASE  
; FILE REFERENCE: 126.70511  
; CURRENT APPLICATION NUMBER: US/09/872,186  
; CURRENT FILING DATE: 2001-06-01  
; PRIOR APPLICATION NUMBER: 09/676,340  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/156,821  
; PRIOR FILING DATE: 1999-09-30  
; PRIOR APPLICATION NUMBER: 60/163,168  
; PRIOR FILING DATE: 1999-11-02  
; PRIOR APPLICATION NUMBER: 60/215,497  
; PRIOR FILING DATE: 2000-06-30  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-872-186-6

Query Match 42.2%; Score 19; DB 5; Length 9;  
Best Local Similarity 80.0%; Pred. No. 1.2e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 ALAF 10  
| : | : |  
DB 5 SLAF 9

RESULT 6  
US-09-856-988B-29  
; Sequence 29, Application US/09856988B  
; GENERAL INFORMATION:  
; APPLICANT: SCHLOM, JEFFREY  
; APPLICANT: HODGE, JAMES  
; APPLICANT: PANICALI, DENNIS  
; TITLE OF INVENTION: A RECOMBINANT VECTOR EXPRESSING MULTIPLE COSTIMULATORY  
; TITLE OF INVENTION: MOLECULES AND USES THEREOF  
; FILE REFERENCE: 38163-0131  
; CURRENT APPLICATION NUMBER: US/09/856,988B  
; CURRENT FILING DATE: 2001-09-24  
; PRIOR APPLICATION NUMBER: PCT/US99/26866  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR APPLICATION NUMBER: 60/111,582  
; PRIOR FILING DATE: 1998-12-09  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 29  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: peptide  
US-09-856-988B-29

Query Match 42.2%; Score 19; DB 5; Length 9;  
Best Local Similarity 80.0%; Pred. No. 1.2e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 ALAFL 10  
:||||  
Db 5 SLAFL 9

## RESULT 7

US-10-647-005-27  
; Sequence 27, Application US/10647005  
; GENERAL INFORMATION:  
; APPLICANT: Cheever, Martin A.  
; Disis, Mary L.  
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN  
; FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
; HER-2/neu ONCOGENE IS ASSOCIATED  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed IP Law Group PLLC  
; STREET: 701 Fifth Avenue Suite 6300  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: US  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/10/647,005  
; APPLICATION NUMBER: US/10/647,005  
; FILING DATE: 21-Aug-2003  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharkey, Richard G.  
; REGISTRATION NUMBER: 32,629  
; REFERENCE/DOCKET NUMBER: 920010.448C10  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: <Unknown>  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 27:  
US-10-647-005-27

Query Match 42.2%; Score 19; DB 6; Length 9;  
Best Local Similarity 80.0%; Pred. No. 1.2e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 ALAFL 10  
:||||  
Db 5 SLAFL 9

## RESULT 8

US-10-657-022-384  
; Sequence 384, Application US/10657022  
; GENERAL INFORMATION:  
; APPLICANT: Simard, John J. L.  
; APPLICANT: Diamond, David C.  
; APPLICANT: Liu, Liping  
; APPLICANT: Liu, Zheng  
; TITLE OF INVENTION: EPITOPE SEQUENCES  
; FILE REFERENCE: MANNK.032A  
; CURRENT APPLICATION NUMBER: US/10/657,022  
; CURRENT FILING DATE: 2003-09-04  
; PRIOR APPLICATION NUMBER: 60/409123  
; PRIOR FILING DATE: 2002-09-06  
; NUMBER OF SEQ ID NOS: 610  
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 384  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homosapiens  
US-10-657-022-384

Query Match 42.2%; Score 19; DB 6; Length 9;  
Best Local Similarity 66.7%; Pred. No. 1.2e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 EALAFL 10  
:||||  
Db 4 ERLAYL 9

## RESULT 9

US-60-485-717-87  
; Sequence 87, Application US/60485717  
; GENERAL INFORMATION:  
; APPLICANT: Bachmann, Martin F  
; APPLICANT: Schwarz, Katrin  
; TITLE OF INVENTION: SYNERGISTIC ENHANCEMENT OF T CELL RESPONSES BY CPGS AND  
; FILE REFERENCE: 1700.0430000  
; CURRENT APPLICATION NUMBER: US/60/485,717  
; CURRENT FILING DATE: 2003-07-10  
; NUMBER OF SEQ ID NOS: 132  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 87  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-60-485-717-87

Query Match 42.2%; Score 19; DB 7; Length 9;  
Best Local Similarity 80.0%; Pred. No. 1.2e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 ALAFL 10  
:||||  
Db 5 SLAFL 9

## RESULT 10

PCT-US03-27706-383  
; Sequence 383, Application PC/TUS0327706  
; GENERAL INFORMATION:  
; APPLICANT: MANNKIND CORPORATION  
; APPLICANT: SIMARD, John J. L.  
; APPLICANT: DIAMOND, David C.  
; APPLICANT: LIU, Liping  
; APPLICANT: LIU, Zheng  
; TITLE OF INVENTION: EPITOPE SEQUENCES  
; FILE REFERENCE: MANNK.032VPC  
; CURRENT APPLICATION NUMBER: PCT/US03/27706  
; CURRENT FILING DATE: 2003-09-05  
; PRIOR APPLICATION NUMBER: US 60/409123  
; PRIOR FILING DATE: 2002-09-06  
; NUMBER OF SEQ ID NOS: 610  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 383  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homosapiens  
PCT-US03-27706-383

Query Match 42.2%; Score 19; DB 1; Length 10;  
Best Local Similarity 66.7%; Pred. No. 9.8e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 EALAFL 10  
:||||  
Db 5 ERLAYL 10

## RESULT 11

PCT-US03-27706-386  
; Sequence 386, Application PC/TUS0327706  
; GENERAL INFORMATION:  
; APPLICANT: MANKIND CORPORATION  
; APPLICANT: SIMARD, John J. L.  
; APPLICANT: DIAMOND, David C.  
; APPLICANT: LIU, Liping  
; APPLICANT: LIU, Zheng  
; TITLE OF INVENTION: EPITOPE SEQUENCES  
; FILE REFERENCE: MANK.032VPC  
; CURRENT APPLICATION NUMBER: PCT/US03/27706  
; CURRENT FILING DATE: 2003-09-05  
; PRIOR APPLICATION NUMBER: US 60/409123  
; PRIOR FILING DATE: 2002-09-06  
; NUMBER OF SEQ ID NOS: 610  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 386  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homosapiens  
PCT-US03-27706-386

Query Match 42.2%; Score 19; DB 1; Length 10;  
Best Local Similarity 66.7%; Pred. No. 9.8e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 EALAF 10  
| | | |  
DB 3 ERLAY 8

## RESULT 12

PCT-US03-27706-387  
; Sequence 387, Application PC/TUS0327706  
; GENERAL INFORMATION:  
; APPLICANT: MANKIND CORPORATION  
; APPLICANT: SIMARD, John J. L.  
; APPLICANT: DIAMOND, David C.  
; APPLICANT: LIU, Liping  
; APPLICANT: LIU, Zheng  
; TITLE OF INVENTION: EPITOPE SEQUENCES  
; FILE REFERENCE: MANK.032VPC  
; CURRENT APPLICATION NUMBER: PCT/US03/27706  
; CURRENT FILING DATE: 2003-09-05  
; PRIOR APPLICATION NUMBER: US 60/409123  
; PRIOR FILING DATE: 2002-09-06  
; NUMBER OF SEQ ID NOS: 610  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 387  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homosapiens  
PCT-US03-27706-387

Query Match 42.2%; Score 19; DB 1; Length 10;  
Best Local Similarity 66.7%; Pred. No. 9.8e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 EALAF 10  
| | | |  
DB 1 ERLAY 6

## RESULT 13

US-09-239-043D-19  
; Sequence 19, Application US/09239043D  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott

; APPLICANT: Vitiello, Maria A.  
; APPLICANT: Livingston, Brian D.  
; APPLICANT: Celis, Esteban  
; APPLICANT: Kubo, Ralph T.  
; APPLICANT: Grey, Howard M.  
; APPLICANT: Chesnut, Robert  
; APPLICANT: EpiImmune Inc.  
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Hepatitis B Virus  
; FILE REFERENCE: 2060.0060007  
; CURRENT APPLICATION NUMBER: US/09/239,043D  
; CURRENT FILING DATE: 1999-01-27  
; PRIOR APPLICATION NUMBER: US 09/189,702  
; PRIOR FILING DATE: 1998-11-10  
; PRIOR APPLICATION NUMBER: US 08/978,291  
; PRIOR FILING DATE: 1997-11-25  
; PRIOR APPLICATION NUMBER: US 08/820,360  
; PRIOR FILING DATE: 1997-03-12  
; PRIOR APPLICATION NUMBER: US 60/013,363  
; PRIOR FILING DATE: 1996-03-13  
; PRIOR APPLICATION NUMBER: US 08/461,603  
; PRIOR FILING DATE: 1995-06-05  
; PRIOR APPLICATION NUMBER: US 08/347,610  
; PRIOR FILING DATE: 1994-12-01  
; PRIOR APPLICATION NUMBER: US 08/344,824  
; PRIOR FILING DATE: 1994-11-23  
; PRIOR APPLICATION NUMBER: US 08/278,634  
; PRIOR FILING DATE: 1994-07-21  
; PRIOR APPLICATION NUMBER: US 08/205,713  
; PRIOR FILING DATE: 1994-03-04  
; PRIOR APPLICATION NUMBER: US 08/197,484  
; PRIOR FILING DATE: 1994-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 2579  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 19  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Orthohepadnaviridae hepatitis B virus  
US-09-239-043D-19

Query Match 42.2%; Score 19; DB 5; Length 10;  
Best Local Similarity 66.7%; Pred. No. 9.8e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 EALAF 10  
| | | |  
DB 1 ELLSFL 6

## RESULT 14

US-09-239-043D-696  
; Sequence 696, Application US/09239043D  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Vitiello, Maria A.  
; APPLICANT: Livingston, Brian D.  
; APPLICANT: Celis, Esteban  
; APPLICANT: Kubo, Ralph T.  
; APPLICANT: Grey, Howard M.  
; APPLICANT: Chesnut, Robert  
; APPLICANT: EpiImmune Inc.  
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Hepatitis B Virus  
; FILE REFERENCE: 2060.0060007  
; CURRENT APPLICATION NUMBER: US/09/239,043D  
; CURRENT FILING DATE: 1999-01-27  
; PRIOR APPLICATION NUMBER: US 09/189,702  
; PRIOR FILING DATE: 1998-11-10  
; PRIOR APPLICATION NUMBER: US 08/978,291  
; PRIOR FILING DATE: 1997-11-25

```
; PRIOR APPLICATION NUMBER: US 08/820,360
; PRIOR FILING DATE: 1997-03-12
; PRIOR APPLICATION NUMBER: US 60/013,363
; PRIOR FILING DATE: 1996-03-13
; PRIOR APPLICATION NUMBER: US 08/461,603
; PRIOR FILING DATE: 1995-06-05
; PRIOR APPLICATION NUMBER: US 08/347,610
; PRIOR FILING DATE: 1994-12-01
; PRIOR APPLICATION NUMBER: US 08/344,824
; PRIOR FILING DATE: 1994-11-23
; PRIOR APPLICATION NUMBER: US 08/278,634
; PRIOR FILING DATE: 1994-07-21
; PRIOR APPLICATION NUMBER: US 08/205,713
; PRIOR FILING DATE: 1994-03-04
; PRIOR APPLICATION NUMBER: US 08/197,484
; PRIOR FILING DATE: 1994-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2579
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 696
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Orthohepadnaviridae hepatitis B virus
US-09-239-043D-696

Query Match          42.2%; Score 19; DB 5; Length 10;
Best Local Similarity 66.7%; Pred. No. 9.8e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      5 EALAF 10
        | | | |
Db       1 ELLSFL 6

RESULT 15
US-09-239-043D-1535
; Sequence 1535, Application US/09239043D
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Vitello, Maria A.
; APPLICANT: Livingston, Brian D.
; APPLICANT: Celis, Esteban
; APPLICANT: Kubo, Ralph T.
; APPLICANT: Grey, Howard M.
; APPLICANT: Chesnut, Robert
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Hepatitis B Virus
; FILE REFERENCE: 2060.0060007
; CURRENT APPLICATION NUMBER: US/09/239,043D
; CURRENT FILING DATE: 1999-01-27
; PRIOR APPLICATION NUMBER: US 09/189,702
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: US 08/978,291
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: US 08/820,360
; PRIOR FILING DATE: 1997-03-12
; PRIOR APPLICATION NUMBER: US 60/013,363
; PRIOR FILING DATE: 1996-03-13
; PRIOR APPLICATION NUMBER: US 08/461,603
; PRIOR FILING DATE: 1995-06-05
; PRIOR APPLICATION NUMBER: US 08/347,610
; PRIOR FILING DATE: 1994-12-01
; PRIOR APPLICATION NUMBER: US 08/344,824
; PRIOR FILING DATE: 1994-11-23
; PRIOR APPLICATION NUMBER: US 08/278,634
; PRIOR FILING DATE: 1994-07-21
; PRIOR APPLICATION NUMBER: US 08/205,713
; PRIOR FILING DATE: 1994-03-04
; PRIOR APPLICATION NUMBER: US 08/197,484
; PRIOR FILING DATE: 1994-02-16
```

```
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2579
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1535
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Orthohepadnaviridae hepatitis B virus
US-09-239-043D-1535

Query Match          42.2%; Score 19; DB 5; Length 10;
Best Local Similarity 66.7%; Pred. No. 9.8e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      5 EALAF 10
        | | | |
Db       1 ELLSFL 6

Search completed: October 7, 2003, 14:42:00
Job time : 14.5 secs
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:31:21 ; Search time 57.5 Seconds  
(without alignments)  
24.844 Million cell updates/sec

Title: US-09-807-512-24  
Perfect score: 43  
Sequence: 1 FLMAQGAML 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues  
Total number of hits satisfying chosen parameters: 179625

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_19Jun03;+  
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	100.0	9	21	AAY70857
2	24	55.8	8	22	ABUS3411
3	24	55.8	9	17	AAW49331
4	24	55.8	9	22	AAW22637
5	24	55.8	9	23	AAU94283
6	24	55.8	9	24	ABR28777
7	24	55.8	9	24	ABR28816
8	24	55.8	9	24	ABR29000
9	24	55.8	9	24	ABR29027
10	24	55.8	9	21	AAY70857
11	24	55.8	9	24	ABR29407
12	24	55.8	9	24	ABR29448
13	24	55.8	9	24	ABR29591
14	23	53.5	8	23	AAU82558
15	23	53.5	8	23	AAU82559
16	23	53.5	8	23	AAU82560
17	22	51.2	7	16	AAU71399
18	22	51.2	7	19	AAW38294
19	22	51.2	7	22	AAW37810
20	22	51.2	8	20	AAW33569
21	22	51.2	9	23	AAU91689
22	22	51.2	9	23	AAU91889
23	22	51.2	9	23	AAU91978
24	22	51.2	9	23	AAU92072
25	21	48.8	8	20	AAW48026
26	21	48.8	8	22	AAW6284
27	21	48.8	8	23	ABJ06986
28	21	48.8	8	23	ABJ07477
29	21	48.8	8	23	ABJ09522
30	21	48.8	8	23	ABJ09823
31	21	48.8	9	15	ABJ73754
32	21	48.8	9	15	AAW73115
33	21	48.8	9	16	AAW78857
34	21	48.8	9	20	AAW47830
35	21	48.8	9	21	AAW73023
36	21	48.8	9	22	AAW32425
37	21	48.8	9	22	AAU02323
38	21	48.8	9	22	AAU02329
39	21	48.8	9	22	AAU02351
40	21	48.8	9	23	ABJ05890
41	21	48.8	9	23	ABJ06039
42	21	48.8	9	23	ABJ07236
43	21	48.8	9	23	ABJ07530
44	21	48.8	9	23	ABJ07962
45	21	48.8	9	23	ABJ08094

ALIGNMENTS

RESULT 1  
AAY70857  
ID AAY70857 standard; peptide; 9 AA.  
XX  
AC AAY70857;  
XX  
DT 31-JUL-2000 (first entry)  
XX  
CAMEL10 Immunogenic peptide of human CAMEL protein.

CAMEL; CTL-recognised Antigen on MELanoma; cytotoxic T lymphocyte; CTL;  
tumour-associated antigen; LAGE-1; NY-ESO-1; anticancer; melanoma;  
human; cancer; immunotherapy; immunogenic peptide; immune response.

OS Homo sapiens.  
XX  
XX  
PN WO200023584-A1.  
XX  
PD 27-APR-2000.  
XX  
PF 15-OCT-1999; 99WO-EP07832.  
XX  
PR 16-OCT-1998; 98EP-0119583.  
XX  
PA (BOEH ) BOEHRINGER INGELHEIM INT GMBH.  
XX (UYHO-) UNIV HOSPITAL LEIDEN.  
PI Schrier PI, Aarnoudse CA, Heider K, Klade C;  
XX WPI; 2000-339685/29.  
XX  
PT Tumor-associated antigen useful for cancer immunotherapy is encoded by

Human cancer-relat  
Human cancer-relat  
Human cancer-relat  
Human cancer-relat  
Llama CDR3 region  
Llama CDR3 region  
Llama CDR3 region  
vpr peptide used t  
Poliovirus amino a  
vpr peptide unique  
Canola CBF1 AP2 ca  
PHORI-All peptide  
PHORI-All peptide  
PHORI-All peptide  
Immunogenic peptid  
Hepatitis B virus  
Hepatitis B virus  
Hepatitis B virus  
Hepatitis B virus  
Antigen fragment 7  
Antigen fragment 2  
HBV env l116-1124  
Immunogenic peptid  
Hepatitis B virus  
HIV peptide SEQ ID  
HLA binding TAGD-1  
HLA binding TAGD-1  
HLA binding TAGD-1  
Hepatitis B virus  
Hepatitis B virus  
Hepatitis B virus  
Hepatitis B virus  
Hepatitis B virus

PT the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA -  
XX  
PS Claim 6; Page 34; 73pp; English.  
XX  
CC The present sequence is an immunogenic peptide CAMEL 10, of the human  
CC tumour-associated antigen CAMEL (cytotoxic T lymphocytes (CTL)-recognised  
CC Antigen on MELANOMA). This peptide has the potential to bind to HLA-A2  
CC and corresponds to residues 10-18 of the CAMEL protein. The CAMEL protein  
CC is encoded by the LAGE-1 gene, a tumour-specific antigen. It is different  
CC from the LAGE-1 protein, since it is translated from a different open  
CC reading frame (ORF-1). It shows strong homology with NY-ESO-1, a melanoma  
CC specific tumour antigen. The tumour-associated antigen displayed on  
CC melanoma cells is recognised by cytotoxic T lymphocytes. CAMEL is  
CC expressed in tumour cell lines, tumour tissues (e.g. breast and lung) and  
CC in restricted number of healthy tissues. This sequence has anticancer  
CC activity. CAMEL tumour antigen and immunogenic peptides derived from it  
CC are useful for cancer immunotherapy. They have the potential to induce an  
CC immune response, by eliciting a CTL response. The DNA molecule is used to  
CC construct recombinant or fusion proteins.  
XX  
SQ Sequence 9 AA;  
Query Match 100.0%; Score 43; DB 21; Length 9;  
Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 FLMAQGAML 9  
Db | | | | | | | | | |  
1 FLMAQGAML 9  
RESULT 2  
ABU53411  
ID ABU53411 standard; Peptide; 8 AA.  
XX  
AC ABU53411;  
XX  
DT 14-APR-2003 (first entry)  
DE Human DNA derived peptide motif #1005.  
XX  
KW Human; gene therapy; vaccine; disease treatment; detection.  
XX  
OS Homo sapiens.  
XX  
PN W0200112659-A2.  
XX  
PD 22-FEB-2001.  
XX  
PF 18-AUG-2000; 2000WO-IB01496.  
XX  
PR 18-AUG-1999; 99US-0149499.  
PR 28-SEP-1999; 99US-0156503.  
XX  
PA (GEHU-) GERMAN HUMAN GENOME PROJECT.  
XX  
PI Wiemann S;  
XX  
DR WPI; 2001-327840/34.  
XX  
PT Nucleic acids having the sequences of clones isolated from libraries of  
PT different human tissues, useful in recombinant DNA methodologies -  
XX  
PS Disclosure; SEQ ID NO 1005; 1095pp; English.  
XX  
CC This invention describes novel polynucleotides and polypeptides isolated  
CC from human cDNA libraries which can be used for gene therapy or in  
CC vaccines. The polynucleotides of the invention and antibodies encoded by  
CC them may be used in the prevention, diagnosis and treatment of diseases  
CC associated with inappropriate polypeptide expression. The products of the  
CC invention may also be used to identify modulators of expression and  
CC activity and to down regulate expression and activity. The antibodies of  
CC the invention may also be used as diagnostic agents for detecting the

CC presence of polypeptides in samples. This sequence represents a peptide  
CC associated with the polypeptides described in the disclosure of the  
CC invention.  
XX  
SQ Sequence 8 AA;  
Query Match 55.8%; Score 24; DB 22; Length 8;  
Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 LMAQGA 6  
Db | | | | |  
1 LMAQGA 5  
RESULT 3  
AAW49331  
ID AAW49331 standard; peptide; 9 AA.  
XX  
AC AAW49331;  
XX  
DT 05-JUN-1998 (first entry)  
XX  
DE Human leucocyte antigen DQ4 binding peptide #222.  
XX  
KW Human leucocyte antigen; HLA-DQ4; combinatorial library;  
KW autoimmune disease; chronic articular rheumatism.  
XX  
OS Synthetic.  
XX  
PN JP08151396-A.  
XX  
PD 11-JUN-1996.  
XX  
PF 28-NOV-1994; 94JP-0292657.  
XX  
PR 28-NOV-1994; 94JP-0292657.  
XX  
PA (TEIJ ) TEIJIN LTD.  
XX  
DR WPI; 1996-329479/33.  
XX  
PT HLA-binding oligopeptide and an immuno:regulator contg it - used in  
PT the treatment of auto:immune disease  
XX  
PS Claim 4; Page 30; 61pp; Japanese.  
XX  
CC This peptide is an example of a peptide which binds to a human leucocyte  
CC antigen HLA-DQ4 molecule. The peptide was isolated from a phagemid  
CC combinatorial library comprising the sequence AAV05953, by screening  
CC with an HLA-DQ4 molecule. The peptide is used for the treatment of  
CC autoimmune disease, or especially for treatment of viral diseases.  
XX  
SQ Sequence 9 AA;  
Query Match 55.8%; Score 24; DB 17; Length 9;  
Best Local Similarity 83.3%; Pred. No. 9.3e+05;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 2 LMAQGA 7  
Db | | | | |  
3 LMAQGA 8  
RESULT 4  
AAW22637  
ID AAW22637 standard; Peptide; 9 AA.  
XX  
AC AAW22637;  
XX  
DT 22-OCT-2001 (first entry)  
XX  
DE HIV peptide SEQ ID NO 522.



```
XX Cytotoxic T cell lymphocyte; CTL; anti-HIV; viral disease; HIV;
KW human immunodeficiency virus; vaccine.
XX
XX Human immunodeficiency virus.
OS Synthetic.
XX WO200155177-A2.
XX
XX 02-AUG-2001.
PD
XX
XX 29-JAN-2001; 2001WO-DK00059.
PF
XX
XX 28-JAN-2000; 2000EP-0610017.
PR
XX 31-JAN-2000; 2000US-0179333.
XX
XX (STAT-) STATENS SERUM INST.
PA
XX
XX Fomsgaard A, Brunak S, Buus S, Corbet S, Lauemoller SL, Hansen J;
PI WPI; 2001-476184/51.
XX
XX The generation of cytotoxic T cell lymphocytes epitopes for use in
PT anti-HIV vaccines.
PT
XX
XX Example 4; Page 66; 383pp; English.
PS
XX
XX The invention relates to identification of cytotoxic T cell lymphocyte
CC (CTL) epitopes (AAM22116-AAM23484) that generate anti-HIV activity. CTL
CC are a major protective mechanism against viral diseases. Antibodies may
CC neutralise extracellular human immunodeficiency virus (HIV) and limit or
CC prevent infection of cells in the host, but CTL will limit viral
CC production by killing the cell. The CTL epitopes are useful in medicine,
CC in the manufacture of vaccines or diagnostic agents.
XX
XX Sequence 9 AA;
SQ
XX
XX Query Match 55.8%; Score 24; DB 22; Length 9;
XX Best Local Similarity 55.6%; Pred. No. 9.3e+05;
XX Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
XX
XX QY 1 FLMAQGAML 9
XX ||| | |
XX 1 FLMIVGGLL 9
XX
XX RESULT 5
XX AAU94283
XX ID AAU94283 standard; Peptide; 9 AA.
XX
XX AC AAU94283;
XX
XX 02-JUL-2002 (first entry)
DT
XX
XX Human novel protein CatrF2E11 HLA binding peptide #116.
DE
XX
XX Human; human leukocyte antigen; HLA; immunogen; 83P2H3; CatrF2E11;
KW calcium transport protein; cancer; prostate cancer; cytostatic;
KW chromosome 7q34; chromosome 12q24.1; T cell; B cell.
XX
XX Homo sapiens.
OS
XX WO200214361-A2.
XX
XX 21-FEB-2002.
PD
XX
XX 17-AUG-2001; 2001WO-US25782.
PF
XX
XX 17-AUG-2000; 2000US-226329P.
PR
XX (AGEN-) AGENSYS INC.
PA
XX
XX Raitano AB, Challita-Eid PM, Faris M, Saffran DC, Afar DEH;
PI
```

```
PI Levin E, Hubert RS, Ge W, Jakobovits A;
XX WPI; 2002-269179/31.
XX
XX Monitoring 83P2H3 gene products for monitoring the presence of cancer
PT in a subject, comprises determining the status of 83P2H3 gene products
PT in a tissue sample from the subject and comparing it to a normal sample
PT
XX
XX Example 11; Page 172; 270pp; English.
PS
XX
XX The invention relates to monitoring 83P2H3 (a calcium transport
CC protein whose gene is located on chromosome 7q34) gene products in a
CC biological sample from a patient who has or is suspected of having
CC cancer (especially prostate cancer), comprises: (a) determining the
CC status of 83P2H3 gene products expressed by cells in a tissue sample from
CC an individual and (b) comparing the status to the status of 83P2H3 gene
CC products in a normal sample. Also included are modulators of 83P2H3
CC function or status, generating antibodies/immune response against
CC 83P2H3 (or related protein CatrF2E11 whose gene is located on chromosome
CC 12q24.1) using identified HLA (human leukocyte antigen) binding
CC peptides derived from the protein, delivering a cytotoxic agent to
CC a cell expressing 83P2H3 by conjugating the agent to an anti-83P2H3
CC antibody, a recombinant protein comprising an antigen-binding region of
CC the antibody, a non-human transgenic animal that produces the recombinant
CC protein, a hybridoma that produces the recombinant protein, a single-
CC chain monoclonal antibody that comprises the variable domains of the
CC heavy and light chains of the anti-83P2H3 antibody, a vector comprising a
CC polynucleotide that encodes the monoclonal antibody and inducing an
CC immune response to a 83P2H3 protein, by providing a 83P2H3-related
CC protein that comprises a T cell or B cell epitope, and contacting the
CC epitope with an immune system T cell or B cell, respectively. The method
CC is useful for monitoring 83P2H3 gene products in a biological sample for
CC monitoring the presence of cancer in an individual. The modulator is
CC useful for inhibiting the growth of cancer cells that express 83P2H3, for
CC treating cancer and the vector is useful for treating a patient with a
CC cancer that expresses 83P2H3. The immunological methods are useful for
CC generating an immune response against 83P2H3, and for detecting the
CC presence of 83P2H3-related protein or polynucleotide in a biological
CC sample from a patient who has or who is suspected of having cancer. The
CC antibody is useful in prostate cancer diagnosis, prognosis, imaging
CC methodologies and treatment, to detect and quantify 83P2H3 and mutant
CC 83P2H3-related proteins, for purifying a 83P2H3-related protein, for
CC isolating 83P2H3 homologues/related molecules, and for generating anti-
CC idiotypic antibodies that mimic the 83P2H3 protein. The present sequence
CC is an HLA binding peptide motif from 83P2H3 or its related protein
CC CatrF2E11.
XX
XX Sequence 9 AA;
SQ
XX
XX Query Match 55.8%; Score 24; DB 23; Length 9;
XX Best Local Similarity 83.3%; Pred. No. 9.3e+05;
XX Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 2 LMAQGA 7
XX | : | | |
XX 2 LVAQGA 7
XX
XX RESULT 6
XX ABR28777
XX ID ABR28777 standard; Peptide; 9 AA.
XX
XX AC ABR28777;
XX
XX 19-MAY-2003 (first entry)
DT
XX
XX Human cancer-related protein 192P2G7 HLA peptide #221.
DE
XX
XX Human; cytostatic; vaccine; cancer; immune response; HLA;
KW human leukocyte antigen.
XX
XX Homo sapiens.
OS
```

```
XX PN WO200283921-A2.
XX PD 24-OCT-2002.
XX PF 10-APR-2002; 2002WO-US11654.
XX PR 10-APR-2001; 2001US-282739P.
XX PR 10-APR-2001; 2001US-283112P.
XX PR 25-APR-2001; 2001US-286630P.
XX PA (AGEN-) AGENSYS INC.
XX PI Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
XX PI Morrison K, Morrison RK, Raitano AB;
XX DR WPI; 2003-075555/07.
XX XX
XX XX New composition comprising a substance that modulates the structure of
XX PT proteins and polynucleotides, useful for therapeutic, prognostic and
XX PT diagnostic reagents for eliciting cellular or humoral immune response
XX PT in cancer patients.
XX PA (AGEN-) AGENSYS INC.
XX PI Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
XX PI Morrison K, Morrison RK, Raitano AB;
XX DR WPI; 2003-075555/07.
XX XX
XX XX The present invention relates to novel human cancer-related genes and
XX PT proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
XX PT proteins are useful for eliciting a humoral or cellular immune response.
XX PT The genes are useful as probes and primers for the amplification and/or
XX PT detection of genes, mRNAs or their fragments, as reagents for the
XX PT diagnosis and/or prognosis of cancer, as coding sequences capable of
XX PT directing the expression of the protein, as tools for modulating or
XX PT inhibiting the expression of genes and/or translation of transcripts, and
XX PT as therapeutic agents. The proteins and peptides are useful as
XX PT therapeutic, prognostic and diagnostic reagents for cancer. The present
XX PT sequence is a human leukocyte antigen (HLA) peptide, used in an example
XX PT from the invention.
XX SQ Sequence 9 AA;
XX Query Match 55.8%; Score 24; DB 24; Length 9;
XX Best Local Similarity 57.1%; Pred. No. 9.3e+05;
XX Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 FLMAQGA 7
XX Db :l::lll
XX 3 YLVSOQA 9
XX XX
XX XX RESULT 7
XX ID ABR28816
XX AC ABR28816 standard; Peptide; 9 AA.
XX XX
XX XX ABR28816;
XX DT 19-MAY-2003 (first entry)
XX DE Human cancer-related protein 192P2G7 HLA peptide #260.
XX XX
XX KW Human; cytostatic; vaccine; cancer; immune response; HLA;
XX KW human leukocyte antigen.
XX OS Homo sapiens.
XX XX
XX PN WO200283921-A2.
XX XX
XX PD 24-OCT-2002.
XX PF 10-APR-2002; 2002WO-US11654.
XX PR 10-APR-2001; 2001US-282739P.
XX PR 10-APR-2001; 2001US-283112P.
XX PR 25-APR-2001; 2001US-286630P.
```

```
XX XX (AGEN-) AGENSYS INC.
XX PA Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
XX PI Morrison K, Morrison RK, Raitano AB;
XX DR WPI; 2003-075555/07.
XX XX
XX XX New composition comprising a substance that modulates the structure of
XX PT proteins and polynucleotides, useful for therapeutic, prognostic and
XX PT diagnostic reagents for eliciting cellular or humoral immune response
XX PT in cancer patients.
XX PA (AGEN-) AGENSYS INC.
XX PI Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
XX PI Morrison K, Morrison RK, Raitano AB;
XX DR WPI; 2003-075555/07.
XX XX
XX XX The present invention relates to novel human cancer-related genes and
XX PT proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
XX PT proteins are useful for eliciting a humoral or cellular immune response.
XX PT The genes are useful as probes and primers for the amplification and/or
XX PT detection of genes, mRNAs or their fragments, as reagents for the
XX PT diagnosis and/or prognosis of cancer, as coding sequences capable of
XX PT directing the expression of the protein, as tools for modulating or
XX PT inhibiting the expression of genes and/or translation of transcripts, and
XX PT as therapeutic agents. The proteins and peptides are useful as
XX PT therapeutic, prognostic and diagnostic reagents for cancer. The present
XX PT sequence is a human leukocyte antigen (HLA) peptide, used in an example
XX PT from the invention.
XX SQ Sequence 9 AA;
XX Query Match 55.8%; Score 24; DB 24; Length 9;
XX Best Local Similarity 57.1%; Pred. No. 9.3e+05;
XX Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 FLMAQGA 7
XX Db :l::lll
XX 1 YLVSOQA 7
XX XX
XX XX RESULT 8
XX ID ABR29000
XX AC ABR29000 standard; Peptide; 9 AA.
XX XX
XX XX ABR29000;
XX DT 19-MAY-2003 (first entry)
XX DE Human cancer-related protein 192P2G7 HLA peptide #444.
XX XX
XX KW Human; cytostatic; vaccine; cancer; immune response; HLA;
XX KW human leukocyte antigen.
XX OS Homo sapiens.
XX XX
XX PN WO200283921-A2.
XX XX
XX PD 24-OCT-2002.
XX PF 10-APR-2002; 2002WO-US11654.
XX PR 10-APR-2001; 2001US-282739P.
XX PR 10-APR-2001; 2001US-283112P.
XX PR 25-APR-2001; 2001US-286630P.
XX XX
XX PA (AGEN-) AGENSYS INC.
XX PI Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
XX PI Morrison K, Morrison RK, Raitano AB;
XX DR WPI; 2003-075555/07.
XX XX
XX XX New composition comprising a substance that modulates the structure of
XX PT proteins and polynucleotides, useful for therapeutic, prognostic and
XX PT diagnostic reagents for eliciting cellular or humoral immune response
XX PT in cancer patients.
```

PT diagnostic reagents for eliciting cellular or humoral immune response  
 XX in cancer patients -  
 PS Claim 13; Page 421; 1021pp; English.  
 XX  
 CC The present invention relates to novel human cancer-related genes and  
 CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and  
 CC proteins are useful for eliciting a humoral or cellular immune response.  
 CC The genes are useful as probes and primers for the amplification and/or  
 CC detection of genes, mRNAs or their fragments, as reagents for the  
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of  
 CC directing the expression of the protein, as tools for modulating or  
 CC inhibiting the expression of genes and/or translation of transcripts, and  
 CC as therapeutic agents. The proteins and peptides are useful as  
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present  
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example  
 CC from the invention.  
 XX  
 SQ Sequence 9 AA;

Query Match 55.8%; Score 24; DB 24; Length 9;  
 Best Local Similarity 57.1%; Pred. No. 9.3e+05;  
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAOGA 7  
 DB 3 YLVSOQA 9  
 :|::|||

RESULT 9  
 ABR29027  
 ID ABR29027 standard; Peptide; 9 AA.  
 XX  
 AC ABR29027;  
 XX  
 DT 19-MAY-2003 (first entry)  
 XX  
 DE Human cancer-related protein 192P2G7 HLA peptide #471.  
 XX  
 KW Human; cytostatic; vaccine; cancer; immune response; HLA;  
 KW human leukocyte antigen.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200283921-A2.  
 XX  
 PD 24-OCT-2002.  
 XX  
 PF 10-APR-2002; 2002WO-US11654.  
 XX  
 PR 10-APR-2001; 2001US-282739P.  
 PR 10-APR-2001; 2001US-283112P.  
 PR 25-APR-2001; 2001US-286630P.  
 XX  
 PA (AGEN-) AGENSYS INC.  
 XX  
 PI Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;  
 PI Morrison K, Morrison RK, Raitano AB;  
 XX  
 WPI; 2003-075555/07.  
 XX  
 CC New composition comprising a substance that modulates the structure of  
 CC proteins and polynucleotides, useful for therapeutic, prognostic and  
 CC diagnostic reagents for eliciting cellular or humoral immune response  
 CC in cancer patients -  
 XX  
 PS Claim 13; Page 422; 1021pp; English.  
 XX  
 CC The present invention relates to novel human cancer-related genes and  
 CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and  
 CC proteins are useful for eliciting a humoral or cellular immune response.  
 CC The genes are useful as probes and primers for the amplification and/or  
 CC detection of genes, mRNAs or their fragments, as reagents for the

CC diagnosis and/or prognosis of cancer, as coding sequences capable of  
 CC directing the expression of the protein, as tools for modulating or  
 CC inhibiting the expression of genes and/or translation of transcripts, and  
 CC as therapeutic agents. The proteins and peptides are useful as  
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present  
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example  
 CC from the invention.  
 XX  
 SQ Sequence 9 AA;

Query Match 55.8%; Score 24; DB 24; Length 9;  
 Best Local Similarity 57.1%; Pred. No. 9.3e+05;  
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAOGA 7  
 DB 1 YLVSOQA 7  
 :|::|||

RESULT 10  
 ABR29186  
 ID ABR29186 standard; Peptide; 9 AA.  
 XX  
 AC ABR29186;  
 XX  
 DT 19-MAY-2003 (first entry)  
 XX  
 DE Human cancer-related protein 192P2G7 HLA peptide #630.  
 XX  
 KW Human; cytostatic; vaccine; cancer; immune response; HLA;  
 KW human leukocyte antigen.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200283921-A2.  
 XX  
 PD 24-OCT-2002.  
 XX  
 PF 10-APR-2002; 2002WO-US11654.  
 XX  
 PR 10-APR-2001; 2001US-282739P.  
 PR 10-APR-2001; 2001US-283112P.  
 PR 25-APR-2001; 2001US-286630P.  
 XX  
 PA (AGEN-) AGENSYS INC.  
 XX  
 PI Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;  
 PI Morrison K, Morrison RK, Raitano AB;  
 XX  
 WPI; 2003-075555/07.  
 XX  
 CC New composition comprising a substance that modulates the structure of  
 CC proteins and polynucleotides, useful for therapeutic, prognostic and  
 CC diagnostic reagents for eliciting cellular or humoral immune response  
 CC in cancer patients -  
 XX  
 PS Claim 13; Page 424; 1021pp; English.  
 XX  
 CC The present invention relates to novel human cancer-related genes and  
 CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and  
 CC proteins are useful for eliciting a humoral or cellular immune response.  
 CC The genes are useful as probes and primers for the amplification and/or  
 CC detection of genes, mRNAs or their fragments, as reagents for the  
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of  
 CC directing the expression of the protein, as tools for modulating or  
 CC inhibiting the expression of genes and/or translation of transcripts, and  
 CC as therapeutic agents. The proteins and peptides are useful as  
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present  
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example  
 CC from the invention.  
 XX  
 SQ Sequence 9 AA;

Query Match 55.8%; Score 24; DB 24; Length 9;  
 Best Local Similarity 57.1%; Pred. No. 9.3e+05;  
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAQGA 7  
 :|::|||  
 Db 3 YLVSQGA 9

RESULT 11

ABR29407  
 ID ABR29407 standard; Peptide; 9 AA.  
 AC ABR29407;  
 XX  
 XX  
 DT 19-MAY-2003 (first entry)  
 XX  
 DE Human cancer-related protein 192P2G7 HLA peptide #851.  
 XX  
 XX Human; cytostatic; vaccine; cancer; immune response; HLA;  
 KW human leukocyte antigen.  
 XX  
 XX Homo sapiens.  
 XX WO200283921-A2.  
 XX 24-OCT-2002.  
 XX 10-APR-2002; 2002WO-US11654.  
 XX 10-APR-2001; 2001US-282739P.  
 PR 10-APR-2001; 2001US-283112P.  
 PR 25-APR-2001; 2001US-286630P.  
 XX (AGEN-) AGENSYS INC.  
 PA Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;  
 PI Morrison K, Morrison RK, Raitano AB;  
 PI WPI; 2003-075555/07.  
 XX New composition comprising a substance that modulates the structure of  
 PT proteins and polynucleotides, useful for therapeutic, prognostic and  
 PT diagnostic reagents for eliciting cellular or humoral immune response  
 PT in cancer patients  
 XX Claim 13; Page 426; 102lpp; English.  
 PS The present invention relates to novel human cancer-related genes and  
 XX proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and  
 CC proteins are useful for eliciting a humoral or cellular immune response.  
 CC The genes are useful as probes and primers for the amplification and/or  
 CC detection of genes, mRNAs or their fragments, as reagents for the  
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of  
 CC directing the expression of the protein, as tools for modulating or  
 CC inhibiting the expression of genes and/or translation of transcripts, and  
 CC as therapeutic agents. The proteins and peptides are useful as  
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present  
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example  
 CC from the invention.  
 XX Sequence 9 AA;  
 SQ Query Match 55.8%; Score 24; DB 24; Length 9;  
 Best Local Similarity 57.1%; Pred. No. 9.3e+05;  
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAQGA 7  
 :|::|||  
 Db 2 YLVSQGA 8

RESULT 12

ABR29407  
 ID ABR29407 standard; Peptide; 9 AA.  
 AC ABR29407;  
 XX  
 XX  
 DT 19-MAY-2003 (first entry)  
 XX  
 DE Human cancer-related protein 192P2G7 HLA peptide #892.  
 XX  
 XX Human; cytostatic; vaccine; cancer; immune response; HLA;  
 KW human leukocyte antigen.  
 XX  
 XX Homo sapiens.  
 XX WO200283921-A2.  
 XX 24-OCT-2002.  
 XX 10-APR-2002; 2002WO-US11654.  
 XX 10-APR-2001; 2001US-282739P.  
 PR 10-APR-2001; 2001US-283112P.  
 PR 25-APR-2001; 2001US-286630P.  
 XX (AGEN-) AGENSYS INC.  
 PA Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;  
 PI Morrison K, Morrison RK, Raitano AB;  
 PI WPI; 2003-075555/07.  
 XX New composition comprising a substance that modulates the structure of  
 PT proteins and polynucleotides, useful for therapeutic, prognostic and  
 PT diagnostic reagents for eliciting cellular or humoral immune response  
 PT in cancer patients  
 XX Claim 13; Page 427; 102lpp; English.  
 PS The present invention relates to novel human cancer-related genes and  
 XX proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and  
 CC proteins are useful for eliciting a humoral or cellular immune response.  
 CC The genes are useful as probes and primers for the amplification and/or  
 CC detection of genes, mRNAs or their fragments, as reagents for the  
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of  
 CC directing the expression of the protein, as tools for modulating or  
 CC inhibiting the expression of genes and/or translation of transcripts, and  
 CC as therapeutic agents. The proteins and peptides are useful as  
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present  
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example  
 CC from the invention.  
 XX Sequence 9 AA;  
 SQ Query Match 55.8%; Score 24; DB 24; Length 9;  
 Best Local Similarity 57.1%; Pred. No. 9.3e+05;  
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAQGA 7  
 :|::|||  
 Db 3 YLVSQGA 9

RESULT 13

ABR29591  
 ID ABR29591 standard; Peptide; 9 AA.  
 AC ABR29591;  
 XX  
 XX 19-MAY-2003 (first entry)  
 DT Human cancer-related protein 192P2G7 HLA peptide #1035.  
 XX Human; cytostatic; vaccine; cancer; immune response; HLA;

KW human leukocyte antigen.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200283921-A2.  
 XX  
 XX 24-OCT-2002.  
 PD  
 XX  
 PF 10-APR-2002; 2002WO-US11654.  
 XX  
 XX 10-APR-2001; 2001US-282739P.  
 PR  
 PR 10-APR-2001; 2001US-283112P.  
 XX  
 XX 25-APR-2001; 2001US-286630P.  
 XX  
 PA (AGEN-) AGENSYS INC.  
 XX  
 XX Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;  
 PI Morrison K, Morrison RK, Raitano AB;  
 XX  
 DR WPI; 2003-075555/07.  
 XX  
 PT New composition comprising a substance that modulates the structure of  
 PT proteins and polynucleotides, useful for therapeutic, prognostic and  
 PT diagnostic reagents for eliciting cellular or humoral immune response  
 PT in cancer patients.  
 XX  
 PS Claim 13; Page 428; 1021pp; English.  
 XX  
 CC The present invention relates to novel human cancer-related genes and  
 CC proteins (AB278120-AB278168 and ABR01789-ABR01861). The genes and  
 CC proteins are useful for eliciting a humoral or cellular immune response.  
 CC The genes are useful as probes and primers for the amplification and/or  
 CC detection of genes, mRNAs or their fragments, as reagents for the  
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of  
 CC directing the expression of the protein, as tools for modulating or  
 CC inhibiting the expression of genes and/or translation of transcripts, and  
 CC as therapeutic agents. The proteins and peptides are useful as  
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present  
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example  
 CC from the invention.  
 XX  
 SQ Sequence 9 AA;  
 Query Match 55.8%; Score 24; DB 24; Length 9;  
 Best Local Similarity 57.1%; Pred. NO. 9.3e+05;  
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 FLMAQGA 7  
 DB :|:|:|  
 3 YLVSQGA 9  
 RESULT 14  
 AAU82558  
 ID AAU82558 standard; peptide; 8 AA.  
 XX  
 AC AAU82558;  
 XX  
 XX 23-APR-2002 (first entry)  
 DT  
 XX  
 DE Llama CDR3 region variable heavy chain fragment #34.  
 KW Llama; phage display library; variable heavy domain fragment; VHH; VH;  
 KW sdAb fragment; single domain anti-idiotypic antibody fragment;  
 KW phage display technology; immune system response; CDR1/H1; CDR2; CDR3;  
 KW complementarity determining region.  
 XX  
 OS Lama guanicoe glama.  
 XX  
 PN WO200190190-A2.  
 XX  
 PD 29-NOV-2001.  
 XX  
 XX Llama; phage display library; variable heavy domain fragment; VHH; VH;  
 KW sdAb fragment; single domain anti-idiotypic antibody fragment;  
 KW phage display technology; immune system response; CDR1/H1; CDR2; CDR3;  
 KW complementarity determining region.  
 XX  
 OS Lama guanicoe glama.  
 XX  
 PN WO200190190-A2.  
 XX  
 XX 29-NOV-2001.  
 DR  
 XX

PF 25-MAY-2001; 2001WO-CA00763.  
 XX  
 PR 26-MAY-2000; 2000US-207234P.  
 XX  
 PA (CANA ) NAT RES COUNCIL CANADA.  
 XX  
 PI Tanha J, Dubuc G, Narang S;  
 XX  
 DR WPI; 2002-083093/11.  
 XX  
 XX New phage display library of variable heavy domain antigen-binding  
 PT fragments derived from llama antibodies, useful for in vitro selection  
 PT against any antigen of interest as a target.  
 XX  
 XX Claim 17; Page 32; 46pp; English.  
 XX  
 CC The present invention relates to a phage display library of variable  
 CC heavy domain (VHH or VH) fragments (sdAb fragments) derived from llama  
 CC antibodies. The library is useful for in vitro selection against any  
 CC antigen of interest as a target. Single domain anti-idiotypic antibody  
 CC fragments are isolated from the library using phage display technology  
 CC and an antibody serving as an antigen. Such anti-idiotypic antibody  
 CC fragments have great potential in evoking the immune system response  
 CC to pathological antigens and in vaccine development. The large size of  
 CC the library considerably increases the probability of isolating from it  
 CC antigen-binding fragments having high affinity to almost any  
 CC predetermined target (antigen of interest). The library eliminates the  
 CC development of anti-idiotypic antibodies by immunisation and allows  
 CC isolation of anti-idiotypic antibodies. AAU82435-AAU82635 represent  
 CC the llama heavy chain domain fragments of the invention.  
 XX  
 SQ Sequence 8 AA;  
 Query Match 53.5%; Score 23; DB 23; Length 8;  
 Best Local Similarity 66.7%; Pred. NO. 9.3e+05;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 MAQQA 8  
 DB :|:|:|  
 1 LSQGA 6  
 RESULT 15  
 AAU82559  
 ID AAU82559 standard; peptide; 8 AA.  
 XX  
 AC AAU82559;  
 XX  
 XX 23-APR-2002 (first entry)  
 DT  
 XX  
 DE Llama CDR3 region variable heavy chain fragment #35.  
 KW Llama; phage display library; variable heavy domain fragment; VHH; VH;  
 KW sdAb fragment; single domain anti-idiotypic antibody fragment;  
 KW phage display technology; immune system response; CDR1/H1; CDR2; CDR3;  
 KW complementarity determining region.  
 XX  
 OS Lama guanicoe glama.  
 XX  
 PN WO200190190-A2.  
 XX  
 PD 29-NOV-2001.  
 XX  
 XX 25-MAY-2001; 2001WO-CA00763.  
 PF  
 PR 26-MAY-2000; 2000US-207234P.  
 XX  
 XX (CANA ) NAT RES COUNCIL CANADA.  
 PA  
 PI Tanha J, Dubuc G, Narang S;  
 XX  
 DR WPI; 2002-083093/11.  
 XX

PT New phage display library of variable heavy domain antigen-binding  
PT fragments derived from llama antibodies, useful for in vitro selection  
PT against any antigen of interest as a target -  
XX  
PS Disclosure; Page 16a; 46pp; English.  
XX

CC The present invention relates to a phage display library of variable  
CC heavy domain (VHH or VH) fragments (sdAb fragments) derived from llama  
CC antibodies. The library is useful for in vitro selection against any  
CC antigen of interest as a target. Single domain anti-idiotypic antibody  
CC fragments are isolated from the library using phage display technology  
CC and an antibody serving as an antigen. Such anti-idiotypic antibody  
CC fragments have great potential in evoking the immune system response  
CC to pathological antigens and in vaccine development. The large size of  
CC the library considerably increases the probability of isolating from it  
CC antigen-binding fragments having high affinity to almost any  
CC predetermined target (antigen of interest). The library eliminates the  
CC development of anti-idiotypic antibodies by immunisation and allows  
CC isolation of anti-idiotypic antibodies. AA082435-AA082635 represent  
CC the llama heavy chain domain fragments of the invention.  
XX

SQ Sequence 8 AA;

Query Match 53.5%; Score 23; DB 23; Length 8;  
Best Local Similarity 66.7%; Pred. No. 9.3e+05;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 MAOGAM 8  
:|||||  
Db 1 LSOGAM 6

Search completed: October 7, 2003, 14:45:32  
Job time : 59.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:42:06 ; Search time 24.5 seconds  
(without alignments)  
35.327 Million cell updates/sec

Title: US-09-807-512-24

Perfect score: 43

Sequence: 1 FLMAQGAML 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 789

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_76:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	37.2	7	2 P00777	NADH2 dehydrogenas
2	15	34.9	5	2 B41225	copper resistance
3	15	34.9	7	2 PT0671	T-cell receptor be
4	15	34.9	8	2 A35180	neutral proteinase
5	14	32.6	6	2 A37765	hypothetical prote
6	14	32.6	6	2 PT0605	T-cell receptor be
7	14	32.6	7	2 A25269	sex pheromone cam3
8	14	32.6	7	2 PT0654	T-cell receptor be
9	14	32.6	8	2 S37141	ipsa protein - Erw
10	14	32.6	9	2 S15850	vitamin D3 26-mono
11	13	30.2	7	2 PT0579	T-cell receptor be
12	13	30.2	7	4 S15597	orf 4 rara 5'-regi
13	12	27.9	5	2 S69237	surface protein te
14	12	27.9	6	2 PT0727	T-cell receptor be
15	12	27.9	6	4 S15596	orf 3 rara 5'-regi
16	12	27.9	7	2 PT0581	T-cell receptor be
17	12	27.9	8	2 PT0527	T-cell receptor be
18	12	27.9	9	1 YFPG	thymic factor - pi
19	12	27.9	9	2 A60957	thymocyte growth p
20	12	27.9	9	2 S07241	litorin - Rohde's
21	12	27.9	9	2 PH1591	Ig H chain V-D-J r
22	11	25.6	3	3 A23751	spinal cord peptid
23	11	25.6	4	1 ECXAA	antho-Rframide neu
24	11	25.6	4	2 A25844	antho-Rf amide neu
25	11	25.6	4	2 PT0271	Ig heavy chain CRD
26	11	25.6	4	2 S09478	globulin IV alpha
27	11	25.6	5	2 A41225	copper resistance
28	11	25.6	6	2 PT0514	T-cell receptor be
29	11	25.6	6	2 PT0593	T-cell receptor be

30	11	25.6	7	2 A44428	platelet aggregati
31	11	25.6	7	2 PT0556	T-cell receptor be
32	11	25.6	7	2 PT0543	T-cell receptor be
33	11	25.6	8	2 A37521	R-phycoerythrin ga
34	11	25.6	8	2 PT0368	Ig gamma chain C r
35	11	25.6	8	2 C39690	neural cell adhesi
36	11	25.6	9	2 A60522	sperm-activating p
37	11	25.6	9	2 C36730	hutu protein - Kle
38	11	25.6	9	2 S13333	alpha/beta-gliadin
39	11	25.6	9	2 S10920	venom protein HR-3
40	11	25.6	9	2 JN0027	[Phe-6]-mosact - S
41	11	25.6	9	2 PT0225	Ig heavy chain CDR
42	11	25.6	9	2 PC2197	zymogen granule me
43	11	25.6	9	2 G85802	hypothetical prote
44	10	23.3	3	3 PT0571	T-cell receptor be
45	10	23.3	4	2 PL0146	carbon monoxide de

ALIGNMENTS

RESULT 1

P00777

NADH2 dehydrogenase (EC 1.6.99.3) 49K chain - fava bean mitochondrion (fragment)  
N:Alternate names: complex I 49K chain; NADH-ubiquinone reductase 49K chain  
C:Species: mitochondrion Vicia faba (fava bean)  
C:Date: 03-May-1994 #sequence\_revision 07-Oct-1994 #text\_change 03-Jun-2002  
C:Accession: P00777

R:Letorme, S.; Boutry, M.

Plant Physiol. 102, 435-443, 1993

A:Title: Purification and preliminary characterization of mitochondrial complex I (NA

A:Reference number: P00775; MUID:94151437; PMID:8108509

A:Accession: P00777

A:Molecule type: protein

A:Residues: 1-7 <LEF>

A:Comment: Complex I, mitochondrial NADH-ubiquinone reductase, is the first of the

ranging from 5K to 75K.

C:Comment: This enzyme catalyzes electron transfer from endogenous NADH to ubiquinone

C:Genetics:

A:Genome: mitochondrion

C:Keywords: electron transfer; mitochondrion; oxidoreductase

Query Match 37.2%; Score 16; DB 2; Length 7;

Best Local Similarity 75.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 QGAM 8

Db 3 QGAV 6

RESULT 2

B41225

copper resistance protein B - Pseudomonas syringae pv. tomato (fragment)

C:Species: Pseudomonas syringae pv. tomato

C:Date: 19-Jun-1992 #sequence\_revision 19-Jun-1992 #text\_change 24-Jun-1993

C:Accession: B41225

R:Cha, J.S.; Cooksey, D.A.

Proc. Natl. Acad. Sci. U.S.A. 88, 8915-8919, 1991

A:Title: Copper resistance in Pseudomonas syringae mediated by periplasmic and outer

A:Reference number: A41225; MUID:92020961; PMID:1924351

A:Accession: B41225

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-5 <CHA>

Query Match 34.9%; Score 15; DB 2; Length 5;

Best Local Similarity 60.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 AOGAM 8

Db 1 AEAAM 5

```
RESULT 3
PT0671
T-cell receptor beta chain V-D-J region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 04-Mar-2000
C:Accession: PT0671; PT0564; PT0537; PT0536; PT0682
J:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0671
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-7 <FEF>
A:Experimental source: day 18 fetal thymus, strain BALB/c, clone 140-1K
A:Accession: PT0537
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-6 <FE2>
A:Experimental source: adult thymus, strain BALB/c, clone 126-1CH
A:Accession: PT0536
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-3 <FE3>
A:Experimental source: adult thymus, strain BALB/c, clone 126-1CC
A:Accession: PT0682
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-3 <FE4>
A:Experimental source: day 18 fetal thymus, strain BALB/c, clone 140-1B
C:Keywords: T-cell receptor

Query Match 34.9%; Score 15; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 QGA 7
Db 5 QGA 7

RESULT 4
A35180
neutral proteinase (EC 3.4.-.-), calcium-activated - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Aug-1990 #sequence_revision 10-Aug-1990 #text_change 21-Mar-1996
C:Accession: A35180
J:Yoshihara, Y.; Ueda, H.; Fujii, N.; Shide, A.; Yajima, H.; Satoh, M.
J. Biol. Chem. 265, 5809-5815, 1990
A:Title: Purification of a novel type of calcium-activated neutral protease from rat brain
A:Reference number: A35180; MUID:90202830; PMID:2318836
A:Accession: A35180
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-8 <YOS>
C:Keywords: hydrolase

Query Match 34.9%; Score 15; DB 2; Length 8;
Best Local Similarity 40.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LMAQG 6
Db 2 LILSEG 6
```

```
RESULT 5
A37765
hypothetical protein (csma 5' region) - Chloroflexus aurantiacus (fragment)
C:Species: Chloroflexus aurantiacus
C:Date: 31-May-1991 #sequence_revision 31-May-1991 #text_change 30-Sep-1993
C:Accession: A37765
J:Theroux, S.J.; Redlinger, T.E.; Fuller, R.C.; Robinson, S.J.
J. Bacteriol. 172, 4497-4504, 1990
A:Title: Gene encoding the 5.7-kilodalton chlorosome protein of Chloroflexus aurantia
A:Reference number: A37765; MUID:90330558; PMID:2376566
A:Accession: A37765
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-6 <THE>
A:Cross-references: GB:M33964

Query Match 32.6%; Score 14; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 MAQ 5
Db 3 MAQ 5

RESULT 6
PT0605
T-cell receptor beta chain V-D-J region (120-1L) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0605
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0605
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-6 <FEF>
A:Experimental source: newborn thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match 32.6%; Score 14; DB 2; Length 6;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 AQGA 7
Db 1 ASGA 4

RESULT 7
A25269
sex pheromone CAM373 - Enterococcus faecalis
N:Alternate names: clumping-inducing agent (CIA)
C:Species: Enterococcus faecalis
C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 18-Jun-1993
C:Accession: A25269
R:Mori, M.; Tanaka, H.; Sakagami, Y.; Isogai, A.; Fujino, M.; Kitada, C.; White, B.A.
FEBS Lett. 206, 69-72, 1986
A:Title: Isolation and structure of the Streptococcus faecalis sex pheromone, CAM373.
A:Reference number: A25269; MUID:87005252; PMID:3093276
A:Accession: A25269
A:Molecule type: protein
A:Residues: 1-7 <MOR>

Query Match 32.6%; Score 14; DB 2; Length 7;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMA 4
Db 3 FLA 6
```



## RESULT 8

PT0654  
T-cell receptor beta chain V-D-J region (121-1BK) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C>Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
C:Accession: PT0654  
R:Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991  
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
A:Reference number: PT0509; MUID:91277601; PMID:1711558  
A:Accession: PT0654  
A>Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-7 <FEE>  
A:Experimental source: day 4 postnatal thymus, strain BALB/c  
C:Keywords: T-cell receptor

Query Match 32.6%; Score 14; DB 2; Length 7;  
Best Local Similarity 75.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 AQAQ 7  
| | |  
DB 1 ASGA 4

## RESULT 9

S37141  
rpsA protein - Erwinia chrysanthemi

C:Species: Erwinia chrysanthemi  
C>Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 08-Oct-1999  
C:Accession: S37141  
R:Douillie, A.; Toussaint, A.; Faelen, M.  
submitted to the EMBL Data Library, August 1993  
A:Description: Identification of the integration host factor genes of E. chrysanthemi.  
A:Reference number: S37139  
A:Accession: S37141  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-8 <DOU>  
A:Cross-references: EMBL:X74750; NID:g399669; PID:CAA52769.1; PID:g581108

Query Match 32.6%; Score 14; DB 2; Length 8;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLMAQG 6  
| | |  
DB 2 FKSAG 7

## RESULT 10

S15850  
vitamin D3 26-monooxygenase (EC 1.14.14.-) cytochrome P450 27 - pig (fragment)  
N:Alternate names: cytochrome P450(26); vitamin D3 26-hydroxylase  
C:Species: Sus scrofa domestica (domestic pig)  
C>Date: 23-Apr-1993 #sequence\_revision 23-Apr-1993 #text\_change 05-Mar-1999  
C:Accession: S15850  
R:Bergman, T.; Postlind, H.

Biochem. J. 276, 427-432, 1991  
A:Title: Characterization of mitochondrial cytochromes P-450 from pig kidney and liver  
A:Reference number: S15850; MUID:91264797; PMID:2049072  
A:Accession: S15850  
A:Molecule type: protein  
A:Residues: 1-9 <BIO>  
C:Genetics:

A:Gene: CYP27  
C:Superfamily: unassigned cytochrome P450; cytochrome P450 homology  
C:Keywords: heme; mitochondrion; monooxygenase; oxidoreductase; transmembrane protein

Query Match 32.6%; Score 14; DB 2; Length 9;

Best Local Similarity 60.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 LMAQG 6  
| | |  
DB 2 LKQG 6

## RESULT 11

PT0579  
T-cell receptor beta chain V-D-J region (141-1IN) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C>Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
C:Accession: PT0579  
R:Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991  
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions  
A:Reference number: PT0509; MUID:91277601; PMID:1711558  
A:Accession: PT0579  
A>Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-7 <FEE>

A:Experimental source: day 19 fetal thymus, strain BALB/c  
C:Keywords: T-cell receptor

Query Match 30.2%; Score 13; DB 2; Length 7;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 MAQG 6  
| | |  
DB 4 LQQG 7

## RESULT 12

S15597  
orf 4 rara 5'-region - human

C:Species: Homo sapiens (man)  
C>Date: 04-Jun-1999 #sequence\_revision 04-Jun-1999 #text\_change 28-Jun-1999  
C:Accession: S15597  
R:Brand, N.J.; Petkovich, M.; Chambon, P.  
Nucleic Acids Res. 18, 6799-6806, 1990  
A:Title: Characterization of a functional promoter for the human retinoic acid receptor  
A:Reference number: S15594; MUID:91088249; PMID:2175878  
A:Accession: S15597  
A:Molecule type: DNA  
A:Residues: 1-7 <BRA>

A:Cross-references: EMBL:X56058; NID:g35876  
A:Note: this ORF from Fig. 2 is not annotated in GenBank entry HSRARA2, release 111.0  
C:Comment: This sequence is not thought to be translated.  
C:Genetics:

A:Gene: GDB:RARA  
A:Cross-references: GDB:120337; OMIM:180240  
A:Map position: 17q12-17q12

Query Match 30.2%; Score 13; DB 4; Length 7;  
Best Local Similarity 28.6%; Pred. No. 2.8e+05;  
Matches 2; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 MAQGAML 9  
| | | : :  
DB 1 MPRGGV 7

## RESULT 13

S69237

-surface protein tetraabrachion heavy chain - Staphylothermus marinus (fragment)  
C:Species: Staphylothermus marinus  
C>Date: 04-Dec-1997 #sequence\_revision 04-Dec-1997 #text\_change 17-Mar-1999  
C:Accession: S69237  
R:Peters, J.; Nitsch, M.; Kuehlmoegen, B.; Golbik, R.; Lupas, A.; Kellermann, J.; Eng  
J. Mol. Biol. 245, 385-401, 1995  
A:Title: Tetraabrachion: a filamentous archaeobacterial surface protein assembly of un

Job time : 24.5 secs

A;Reference number: S69237; MUID:95139068; PMID:7837271  
A;Accession: S69237  
A;Molecule type: protein  
A;Residues: 1-5 <PEP>  
A;Experimental source: strain F1, DSM 3639  
C;Keywords: cell wall; glycoprotein; heat-stable protein

Query Match 27.9%; Score 12; DB 2; Length 5;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 6 GAML 9  
| : |  
Db 1 GTLL 4

## RESULT 14

PT0727  
T-cell receptor beta chain V-D-J region (161-2F) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
C;Accession: PT0727  
R;Peeney, A.J.  
J. Exp. Med. 174, 115-124, 1991  
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
A;Reference number: PT0509; MUID:91277601; PMID:1711558  
A;Accession: PT0727  
A;Status: translation not shown  
A;Molecule type: DNA  
A;Residues: 1-6 <FE>  
A;Experimental source: newborn thymus, strain BALB/c  
C;Keywords: T-cell receptor

Query Match 27.9%; Score 12; DB 2; Length 6;  
Best Local Similarity 40.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 AOGAM 8  
| : |  
Db 1 ASGGL 5

## RESULT 15

SL5596  
orf 3 rara 5'-region - human  
C;Species: Homo sapiens (man)  
C;Date: 04-Jun-1999 #sequence\_revision 04-Jun-1999 #text\_change 28-Jun-1999  
C;Accession: SL5596  
R;Brand, N.J.; Petkovich, M.; Chambon, P.  
Nucleic Acids Res. 18, 6799-6806, 1990  
A;Title: Characterization of a functional promoter for the human retinoic acid receptor-  
A;Reference number: SL5594; MUID:91088249; PMID:2175878  
A;Accession: SL5596  
A;Molecule type: DNA  
A;Residues: 1-6 <BRA>  
A;Cross-references: EMBL:X56058; NID:g35876  
A;Note: this ORF from Fig. 2 is not annotated in GenBank entry HSRARA2, release 111.0  
C;Comment: This sequence is not thought to be translated.  
C;Genetics:  
A;Gene: GDB:RARA  
A;Cross-references: GDB:120337; OMIM:180240  
A;Map position: 17q12-17q12

Query Match 27.9%; Score 12; DB 4; Length 6;  
Best Local Similarity 60.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 MAQGA 7  
| : |  
Db 1 MAPSA 5

Search completed: October 7, 2003, 14:49:04

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:32:11 ; Search time 13 Seconds  
(without alignments)  
32.557 Million cell updates/sec

Title: US-09-807-512-24  
Perfect score: 43  
Sequence: 1 FILMAGGAML 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 251

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	14	32.6	7	CIA_ENTFA	P11932 enterococcu
2	14	32.6	8	RS1_ERWCH	P37985 erwinia chr
3	13	30.2	8	FAR3_PENMO	P83318 penaeus mon
4	12	27.9	8	WPI_PERAT	P83195 perkinsus a
5	12	27.9	9	LITR_PHYRO	P08946 phyllomedus
6	12	27.9	9	PGUR_DIAAB	P81179 diaprepes a
7	12	27.9	9	THVE_PIG	P01255 sus scrofa
8	11	25.6	7	UH11_RAT	P56576 rattus norv
9	11	25.6	8	ALL5_CYDPO	P82156 cydia pomon
10	11	25.6	8	COM2_CONPU	P58785 conus purpu
11	11	25.6	8	CPD1_ENTFA	P13269 enterococcu
12	11	25.6	8	LCK3_LEUMA	P21142 leucophaea
13	11	25.6	9	HUTU_KLEAE	P12381 klebsiella
14	11	25.6	9	MOSF_CLYJA	P19853 clypeaster
15	11	25.6	9	OXYT_RAJCL	P42994 raja clav
16	10	23.3	4	DCMS_PSECH	P19918 pseudomonas
17	10	23.3	4	FLRF_HIRME	P42561 hirudo medi
18	10	23.3	4	FLRN_ATEL	P58707 anthopleura
19	10	23.3	5	BI0B_CITFR	P12997 citrobacter
20	10	23.3	6	CIP2_MYED	P13737 mytilus edu
21	10	23.3	7	FAR1_HELTI	P41871 helisoma tr
22	10	23.3	7	FAR1_MACRS	P83274 macrobrachi
23	10	23.3	7	FAR1_PROCL	P38499 procambaru
24	10	23.3	7	FAR2_PROCL	P38498 procambaru
25	10	23.3	7	PPH2_LYCES	P83379 lycopersico
26	10	23.3	8	CLP_THICU	P80488 thiobacillu
27	10	23.3	8	FAR1_PANRE	P41872 panagrellus
28	10	23.3	8	FAR1_PENMO	P83316 penaeus mon
29	10	23.3	8	FAR2_MACRS	P83275 macrobrachi
30	10	23.3	8	FAR3_HOMAM	P41486 homarus ame
31	10	23.3	8	FAR4_HOMAM	P41487 homarus ame
32	10	23.3	8	FAR8_CALVO	P41863 calliphora
33	10	23.3	8	GLUR_HUMAN	P02729 homo sapien

34 10 23.3 8 1 LCR8\_LEUMA P19990 leucophaea  
35 10 23.3 8 1 NPB\_BOVIN P15507 bos taurus  
36 10 23.3 8 1 UF06\_MOUSE P38644 mus musculus  
37 10 23.3 9 1 DSIP\_RABIT P01158 oryctolagus  
38 10 23.3 9 1 FAR2\_PANRE P41873 panagrellus  
39 10 23.3 9 1 FAR3\_MACRS P83276 macrobrachi  
40 10 23.3 9 1 FAR4\_PENMO P83319 penaeus mon  
41 10 23.3 9 1 FAR6\_CALVO P41861 calliphora  
42 10 23.3 9 1 FAR6\_MACRS P83279 macrobrachi  
43 10 23.3 9 1 FAR8\_MACRS P83281 macrobrachi  
44 10 23.3 9 1 FARP\_CALSI P38495 callinectes  
45 10 23.3 9 1 NEUU\_CAVPO P34966 cavia porce

## ALIGNMENTS

RESULT 1  
CIA\_ENTFA STANDARD; PRT; 7 AA.  
ID P11932;  
AC 01-OCT-1989 (Rel. 12, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 01-FEB-1991 (Rel. 17, Last annotation update)  
DE Sex pheromone CAM373 (Clumping-inducing agent) (CIA).  
OS Enterococcus faecalis (Streptococcus faecalis).  
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.  
OX NCBI\_TaxID=1351;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=87005252; PubMed=3093276;  
RA Mori M., Tanaka H., Sakagami Y., Isogai A., Fujino M., Kitada C.,  
White B.A., An F.Y., Clewell D.B., Suzuki A.;  
RT "Isolation and structure of the Streptococcus faecalis sex pheromone,  
CAM373.";  
RL FEBS Lett. 206:69-72(1986).  
CC -!- FUNCTION: CAM373 INDUCES MATING RESPONSE OF DONOR CELLS  
HARBORING PAM373.  
CC -!- MISCELLANEOUS: THE N-TERMINUS IS POSSIBLY RESPONSIBLE FOR  
SPECIFICITY OF PHEROMONES TO PLASMIDS.  
CC -!- SIMILARITY: C-TERMINAL TO THE PHEROMONES CPD1 AND CAD1.  
DR PIR; A25269; A25269.  
KW Pheromone.  
SQ SEQUENCE 7 AA; 734 MW; 75BDD72059C05DB0 CRC64;  
  
Query Match 32.6%; Score 14; DB 1; Length 7;  
Best Local Similarity 50.0%; Pred. No. 1.3e+05;  
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMA 4  
Db 3 FILA 6

RESULT 2  
RS1\_ERWCH STANDARD; PRT; 8 AA.  
ID P37985;  
AC 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE 30S ribosomal protein S1 (Fragment).  
GN RPSA.  
OS Erwinia chrysanthemi.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Pectobacterium.  
OX NCBI\_TaxID=556;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=3937;  
RA Doullie A., Toussaint A., Faelen M.;  
RL Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: BINDS MRNA; THUS FACILITATING RECOGNITION OF THE

```

CC INITIATION POINT. IT IS NEEDED TO TRANSLATE MRNA WITH A SHORT
CC SHINE-DALGARNO (SD) PURINE-RICH SEQUENCE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE S1P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X74750; CAA52769.1; -.
DR PIR; S37141.
KW Ribosomal protein; Repeat; RNA-binding.
FT NON_TER 1
SQ SEQUENCE 8 AA; 837 MW; 9E18733DC5B339CD CRC64;

Query Match 32.6%; Score 14; DB 1; Length 8;
Best Local Similarity 50.0%; Pred. No. 1.3e+05;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLMAQG 6
DB 2 FKSAKG 7

RESULT 3
FAR3_PENMO
ID FAR3_PENMO STANDARD; PRT; 9 AA.
AC P83318;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE FMRamide-like neuropeptide FLP3 (AQPSMRLRF-amide).
OS Penaeus monodon (Penaeid shrimp).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;
OC Penaeidae; Penaeus.
OX NCBI_TaxID=6687;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE=Eyestalk;
RX MEDLINE=21956277; PubMed=11959015;
RA Sithigorngul P., Pupum J., Krungkasem C., Longvatt S.,
RA Chaivuthangkura P., Sithigorngul W., Petsom A.;
RT "Seven novel FMRamide-like neuropeptide sequences from the eyestalk
RT of the giant tiger prawn Penaeus monodon.";
RL Comp. Biochem. Physiol. 131B:325-337(2002).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MASS SPECTROMETRY: MW=1121.2; METHOD=MALDI.
CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 9
SQ SEQUENCE 9 AA; 1105 MW; CCD107340685A776 CRC64;

Query Match 30.2%; Score 13; DB 1; Length 9;
Best Local Similarity 60.0%; Pred. No. 1.3e+05;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 AQCAM 8
DB 1 AQPSM 5

RESULT 4
WPI_PERAT
ID WPI_PERAT STANDARD; PRT; 8 AA.
AC P83195;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

```

DE Wall protein-1 (PWP-1) (Fragment).
OS Perkinsus atlanticus.
OC Eukaryota; Alveolata; Perkinsia; Perkinsidae; Perkinsida; Perkinsida; Perkinsus.
OX NCBI_TaxID=106964;
RN [1]
RP SEQUENCE, FUNCTION, SUBCELLULAR LOCATION, AND DEVELOPMENTAL STAGE.
RX MEDLINE=22044350; PubMed=12049410;
RA Montes J.F., Dufort M., Llado A., Garcia-Valero J.;
RT "Characterization and immunolocalization of a main proteinaceous
RT component of the cell wall of the protozoan parasite Perkinsus
RT atlanticus.";
RL Parasitology 124:477-484(2002).
CC -1- FUNCTION: Is a major protein component of the cell wall. May play
CC a key role in the organization of the cell wall and in promoting
CC the survival of this parasite.
CC -1- SUBCELLULAR LOCATION: Cell wall. Disulfide-linked to other cell
CC wall components.
CC -1- DEVELOPMENTAL STAGE: Expressed throughout all walled developmental
CC stages.
KW Cell wall.
FT NON_TER 8
SQ SEQUENCE 8 AA; 765 MW; F1787DD87B1AAB16 CRC64;

Query Match 27.9%; Score 12; DB 1; Length 8;
Best Local Similarity 66.7%; Pred. No. 1.3e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 OGA 7
DB 4 EGA 6

RESULT 5
LITR_PHYRO
ID LITR_PHYRO STANDARD; PRT; 9 AA.
AC P08946;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Rhodei-litorin.
OS Phyllomedusa rohdei (Rhode's leaf frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Phyllomedusinae; Phyllomedusa.
OX NCBI_TaxID=8394;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin secretion;
RX MEDLINE=85127560; PubMed=3838283;
RA Barra D., Erspamer G.F., Simmaco M., Bossa F., Melchiorri P.,
RA Erspamer V.;
RT "Rhodei-litorin: a new peptide from the skin of Phyllomedusa rohdei.";
RL FEBS Lett. 182:53-56(1985).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Skin.
CC -1- SIMILARITY: BELONGS TO THE BOMBESIN/NEUROMEDIN B/RANATENSIN
CC FAMILY.
DR PIR; S07241; S07241.
DR InterPro: IPR000874; Bombesin.
DR PFAM; PF02044; Bombesin; 1.
DR PROSITE; PS00257; BOMBESIN; 1.
KW Amphibian defense peptide; Bombesin family; Amidation;
KW Pyrrolidone carboxylic acid.
FT MOD_RES 1
FT MOD_RES 9
SQ SEQUENCE 9 AA; 1090 MW; 4ECC1B861ADC377 CRC64;

Query Match 27.9%; Score 12; DB 1; Length 9;
Best Local Similarity 60.0%; Pred. No. 1.3e+05;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 LMAQG 6
DB 1

```

Db 2 LWATG 6

RESULT 6  
PGLR\_DIAAB  
ID PGLR\_DIAAB STANDARD; PRT; 9 AA.  
AC P81179;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Endo-polygalacturonase (PG) (EC 3.2.1.15) (Fragment).  
OS Diaprepes abbreviatus (Sugarcane rootstalk borer weevil).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;  
OC Phycophaga; Curculionidae; Entiminae; Entimini; Diaprepes.  
OX NCBI\_TaxID=13040;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Larval gut;  
RA Doostdar H., McCollum T.G., Mayer R.T.;  
RT "Purification and characterization of an endo-polygalacturonase from the gut of West Indies sugarcane rootstalk borer weevil (Diaprepes abbreviatus L.) larvae."  
RL Comp. Biochem. Physiol. 118B:861-867(1997).  
CC -!- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-alpha-D-galactosiduronic linkages in pectate and other galacturonans.  
CC -!- INDUCTION: INHIBITED BY CITRUS PGIP.  
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN IS: 9.4, ITS MW IS: 44.5 kDa.  
CC -!- SIMILARITY: WEAK, TO OTHER POLYGALACTURONASES.  
KW Hydrolase; Glycosidase; Cell wall.  
FT NON\_TER 9  
SQ SEQUENCE 9 AA; 1041 MW; 1F49087042DB41BB CRC64;

Query Match 27.9%; Score 12; DB 1; Length 9;  
Best Local Similarity 16.7%; Pred. No. 1.3e+05;  
Matches 1; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLMAQG 6  
Db 4 YVIGRG 9

RESULT 7  
THYF\_PIG  
ID THYF\_PIG STANDARD; PRT; 9 AA.  
AC P01255;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Thymic factor.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=78026571; PubMed=914862;  
RA Pleau J.-M., Dardenne M., Blouquit Y., Bach J.-F.;  
RT "Structural study of circulating thymic factor: a peptide isolated from pig serum. II. Amino acid sequence."  
RL J. Biol. Chem. 252:8045-8047(1977).  
CC -!- MISCELLANEOUS: THE BIOLOGICAL SOURCE(S) AND PHYSIOLOGICAL ACTIVITIES OF THIS PEPTIDE HAVE NOT BEEN DETERMINED.  
CC PIR: A01523; YFPG.  
DR Pyrrolidone carboxylic acid.  
KW Pyrrolidone carboxylic acid.  
FT MOD\_RES 1  
SQ SEQUENCE 9 AA; 876 MW; D500B87866C5B33D CRC64;

Query Match 27.9%; Score 12; DB 1; Length 9;  
Best Local Similarity 66.7%; Pred. No. 1.3e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 AOG 6  
Db 4 SOG 6

RESULT 8  
UHIL\_RAT  
ID UHIL\_RAT STANDARD; PRT; 7 AA.  
AC P56576;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE Unknown protein from 2D-page of heart tissue (Spot P11) (Fragment).  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=Wistar; TISSUE=Heart;  
RA Li X.-P., Pleissner K.-P., Scheler C., Regitz-Zagrosek V., Salikov J., Jungblut P.R.;  
RL Submitted (SEP-1998) to the SWISS-PROT data bank.  
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN PROTEIN IS: 8.5, ITS MW IS: 42 kDa.  
CC UNSURE 2 2 OR A.  
FT NON\_TER 7  
SQ SEQUENCE 7 AA; 775 MW; 6866DB040DC5A6B0 CRC64;

Query Match 25.6%; Score 11; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 QG 6  
Db 6 QG 7

RESULT 9  
ALL5\_CVDPO  
ID ALL5\_CVDPO STANDARD; PRT; 8 AA.  
AC P82156;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Cydiastatin 5.  
OS Cydia pomonella (Codling moth).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  
OC Tortricidae; Tortricidae; Olethreutinae; Cydia.  
OX NCBI\_TaxID=82600;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Larva;  
RX MEDLINE=98054539; PubMed=9392829;  
RA Duve H., Johnsen A.H., Maestri J.-L., Scott A.G., Winstanley D., Davey M., East P.D., Thorpe A.;  
RT "Lepidopteran peptides of the allatostatin superfamily."  
RL Peptides 18:1301-1309(1997).  
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
KW Neuropeptide; Amidation.  
FT MOD\_RES 8 8  
SQ SEQUENCE 8 AA; 898 MW; 922879CABB58640D CRC64;

Query Match 25.6%; Score 11; DB 1; Length 8;  
Best Local Similarity 66.7%; Pred. No. 1.3e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 AOG 6  
Db 1 ARG 3

## RESULT 10

COW2\_CONPU  
ID COW2\_CONPU STANDARD; PRT; 8 AA.  
AC P58785;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Leu-contryphan-P.  
OS Conus purpurascens (Purple cone).  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;  
OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;  
OC Neogastropoda; Conoidea; Conidae; Conus.  
OX NCBI\_TaxID=41690;  
RN [1]  
RP SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.  
RC STRAIN=Clipperton Island; TISSUE=Venom;  
RX MEDLINE=99388839; PubMed=10461743;  
RA Jacobsen R.B., Jimenez E.C., De la Cruz R.G.C., Gray W.R., Cruz L.J.,  
RA Olivera B.M.;  
RT "A novel D-leucine-containing Conus peptide: diverse conformational  
RT dynamics in the contryphan family.";  
RL J. Pept. Res. 54:93-99(1999).  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- TISSUE SPECIFICITY: Expressed by the venom duct.  
CC -1- MASS SPECTROMETRY: MW=888.4; METHOD=LSIMS.  
CC -1- SIMILARITY: BELONGS TO THE CONTRYPHAN FAMILY.  
KW Toxin; Hydroxylation; D-amino acid.  
FT DISULFID 2 4  
FT MOD\_RES 4 4 D-LEUCINE.  
SQ SEQUENCE 8 AA; 890 MW; 75A367672732CEB8 CRC64;

Query Match 25.6%; Score 11; DB 1; Length 8;  
Best Local Similarity 50.0%; Pred. No. 1.3e+05;  
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 GAML 9  
I : I  
Db 1 GCVL 4

## RESULT 11

CPD1\_ENTFA  
ID CPD1\_ENTFA STANDARD; PRT; 8 AA.  
AC P13269;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 01-FEB-1991 (Rel. 17, Last annotation update)  
DE Sex pheromone cPD1  
OS Enterococcus faecalis (Streptococcus faecalis).  
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.  
OX NCBI\_TaxID=1351;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=85040388; PubMed=6436978;  
RA Suzuki A., Mori M., Sagakami Y., Isogai A., Fujino M., Kitada C.,  
RA Craig R.A., Clewell D.B.;  
RT "Isolation and structure of bacterial sex pheromone, cPD1.";  
RL Science 226:849-850(1984).  
CC -1- FUNCTION: cPD1 IS INVOLVED IN THE CONJUGATIVE TRANSFER OF THE  
CC BACTERIOCIN PLASMID PPD1.  
KW Pheromone.  
SQ SEQUENCE 8 AA; 913 MW; 8665B729C682C729 CRC64;

Query Match 25.6%; Score 11; DB 1; Length 8;  
Best Local Similarity 66.7%; Pred. No. 1.3e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLM 3  
I : I  
Db 1 FLV 3

## RESULT 12

LOCK3\_LEUMA  
ID LOCK3\_LEUMA STANDARD; PRT; 8 AA.  
AC P21142;  
DT 01-MAY-1991 (Rel. 18, Created)  
DT 01-MAY-1991 (Rel. 18, Last sequence update)  
DT 01-MAY-1991 (Rel. 18, Last annotation update)  
DE Leucokinin III (L-III).  
OS Leucophaea maderae (Madeira cockroach).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blaberoidea;  
OC Blaberidae; Leucophaea.  
OX NCBI\_TaxID=6988;  
RN [1]  
RP SEQUENCE, AND SYNTHESIS.  
RC TISSUE=Head;  
RA Holman G.M., Cook B.J., Nachman R.J.;  
RT "Primary structure and synthesis of two additional neuropeptides  
RT from Leucophaea maderae: members of a new family of  
RT Cephalomyotropins.";  
RL Comp. Biochem. Physiol. 84C:271-276(1986).  
CC -1- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE  
CC ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).  
CC -1- SIMILARITY: TO THE OTHER LEUCOKININS.  
KW Neuropeptide; Amidation.  
FT MOD\_RES 8 8 AMIDATION.  
SQ SEQUENCE 8 AA; 910 MW; DC6365B449C866DA CRC64;

Query Match 25.6%; Score 11; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 QG 6  
I : I  
Db 2 QG 3

## RESULT 13

HUTU\_KLEAE  
ID HUTU\_KLEAE STANDARD; PRT; 9 AA.  
AC P12351;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Urocanate hydratase (EC 4.2.1.49) (Urocanase) (Imidazolonepropionate  
DE hydrolase) (Fragment).  
GN HUTU.  
OS Klebsiella aerogenes.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Klebsiella.  
OX NCBI\_TaxID=28451;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88198018; PubMed=2834335;  
RA Nieuwkoop A.J., Baldauf S.A., Hudspeth M.E.S., Bender R.A.;  
RT "Bidirectional promoter in the hut(P) region of the histidine  
RT utilization (hut) operons from Klebsiella aerogenes.";  
RL J. Bacteriol. 170:2240-2246(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90368611; PubMed=2203754;  
RA Schwacha A., Bender R.A.;  
RT "Nucleotide sequence of the gene encoding the repressor for the  
RT histidine utilization genes of Klebsiella aerogenes.";  
RL J. Bacteriol. 172:5477-5481(1990).  
CC -1- CATALYTIC ACTIVITY: 3-(5-oxo-4, 5-dihydro-3-H-imidazol-4-  
CC yl)propanoate + urocanate + H(2)O.  
CC -1- COFACTOR: Binds 1 NAD per subunit (By similarity).  
CC -1- PATHWAY: Histidine degradation; second step.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -1- SIMILARITY: BELONGS TO THE UROCANASE FAMILY.  
-----

This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----

DR EMBL; M19665; AAA25078.1; -;  
 DR EMBL; M34604; AAA25076.1; -;  
 DR HAMAP; MF\_00577; -; 1.  
 DR InterPro; IPR000193; Urocanase.  
 DR PROSITE; PS01233; UROCANASE; PARTIAL.  
 KW Histidine metabolism; Lyase; NAD.  
 FT NON\_TER 9  
 SQ SEQUENCE 9 AA; 1140 MW; 970FC41B5325A6C5 CRC64;

Query Match 25.6%; Score 11; DB 1; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 1.3e+05;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 MAQ 5  
 Db 1 MSQ 3

## RESULT 14

MOSF\_CLYJA STANDARD; PRT; 9 AA.  
 AC P19853;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 01-FEB-1991 (Rel. 17, Last annotation update)  
 DE [Phe-6]-mosact.  
 OS Clypeaster japonicus (Sand dollar).  
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
 OC Echinoidea; Euechinoidea; Gnathostomata; Clypeasteroidea;  
 OC Clypeasteridae; Clypeaster.  
 OX NCBI\_TaxID=7644;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Egg jelly;  
 RA Suzuki N, Kurita M, Yoshino K.I., Kajiuira H., Nomura K.,  
 RA Yanaguchi M.;  
 RT "Purification and structure of mosact and its derivatives from the  
 RT egg jelly of the sea urchin Clypeaster japonicus.";  
 RL Zool. Sci. 4:649-656(1987).  
 CC -!- FUNCTION: Stimulates sperm respiration and motility.  
 DR PIR: JN0027; JN0027  
 SQ SEQUENCE 9 AA; 924 MW; 93245729CDC5BAB5 CRC64;

Query Match 25.6%; Score 11; DB 1; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 1.3e+05;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLM 3  
 Db 6 FLI 8

## RESULT 15

OXYT\_RAJCL STANDARD; PRT; 9 AA.  
 AC P42994;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE Glumitocin.  
 OS Raja clavata (Thornback ray).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
 OC Elasmobranchii; Squalia; Hypnosqualea; Pristiorajae; Batoidea;  
 OC Rajiformes; Rajidae; Raja.  
 OX NCBI\_TaxID=7781;  
 RN [1]  
 RP SEQUENCE.

RX MEDLINE=66123415; PubMed=5880565;  
 RA Acher R., Chauvet J., Chauvet M.-T., Crepy D.;  
 RT "Phylogeny of neurohypophyseal peptides: isolation of a new hormone,  
 RT glumitocin (Ser 4-Gln 8-ocytocin) present in a cartilaginous fish,  
 RT the ray (Raja clavata).";  
 RL Biochim. Biophys. Acta 107:393-396(1965).  
 CC -!- FUNCTION: ANTIDIURETIC HORMONE.  
 CC -!- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.  
 DR InterPro; IPR000981; Neurhyp\_horm.  
 DR Pfam; PF00220; hormone4; 1  
 DR PROSITE; PS00264; NEUROHYPOPHYS\_HORM; 1.  
 KW Hormone; Amidation.  
 FT DISULFID 1 6  
 FT MOD\_RES 9 9  
 SQ SEQUENCE 9 AA; 984 MW; 17E9C76EB45B04B CRC64;

Query Match 25.6%; Score 11; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 QG 6  
 Db 8 QG 9

Search completed: October 7, 2003, 14:46:05  
 Job time : 14 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:41:36 ; Search time 58.5 seconds  
(without alignments)  
39.700 Million cell updates/sec

Title: US-09-807-512-24  
Perfect score: 43  
Sequence: 1 FLMAQGAML 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues  
Total number of hits satisfying chosen parameters: 775

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_23.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvivirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16	37.2	8	13 Q9PS69	Q9PS69 gallus gall
2	15	34.9	9	4 Q9BQT4	Q9BQT4 homo sapien
3	14	32.6	9	4 Q9NYH5	Q9NYH5 homo sapien
4	14	32.6	9	5 Q9TWD6	Q9TWD6 leptinotars
5	13	30.2	8	2 P83152	P83152 anabaena sp
6	13	30.2	8	11 Q8R4D8	Q8R4D8 mus musculu
7	13	30.2	8	12 Q83977	Q83977 influenzavi
8	13	30.2	8	12 Q83349	Q83349 murine coro
9	13	30.2	9	2 Q43960	Q43960 azotobacter
10	13	30.2	9	2 P72149	P72149 pseudomonas
11	13	30.2	9	4 Q95953	Q95953 homo sapien
12	13	30.2	9	8 Q9TLM7	Q9TLM7 laurancia v
13	13	30.2	9	15 Q85599	Q85599 moloney mur
14	12	27.9	8	2 Q93SP2	Q93SP2 pseudomonas
15	12	27.9	8	4 Q9UCN4	Q9UCN4 homo sapien
16	12	27.9	8	8 Q94VC1	Q94VC1 varanus rud

17	12	27.9	8	8 Q94VA0	Q94VA0 varanus sem
18	12	27.9	8	8 Q8HR64	Q8HR64 clivia caul
19	12	27.9	8	10 Q9SAV7	Q9SAV7 dioscorea t
20	12	27.9	8	12 Q89498	Q89498 murine hepa
21	12	27.9	9	4 Q16386	Q16386 homo sapien
22	12	27.9	9	8 Q9MMF4	Q9MMF4 buteo rufin
23	12	27.9	9	8 Q9MMG9	Q9MMG9 buteo buteo
24	11	25.6	8	2 Q9X3K1	Q9X3K1 prochloroco
25	11	25.6	8	2 Q9R4M3	Q9R4M3 enterococcu
26	11	25.6	8	4 Q60773	Q60773 homo sapien
27	11	25.6	8	6 Q28866	Q28866 megaloptera n
28	11	25.6	8	13 Q90ZV5	Q90ZV5 fulica leuc
29	11	25.6	9	2 Q46179	Q46179 clostridium
30	11	25.6	9	4 Q16605	Q16605 homo sapien
31	11	25.6	9	4 Q9UK44	Q9UK44 homo sapien
32	11	25.6	9	5 Q9TAX7	Q9TAX7 manduca sex
33	11	25.6	9	7 Q78225	Q78225 mus musculu
34	11	25.6	9	7 Q78226	Q78226 mus musculu
35	11	25.6	9	12 Q69100	Q69100 herpes simp
36	11	25.6	9	13 Q8AUM7	Q8AUM7 carassius a
37	11	25.6	9	15 Q85723	Q85723 simian sarc
38	11	25.6	9	16 Q8X4G1	Q8X4G1 escherichia
39	10	23.3	7	2 Q8KMS3	Q8KMS3 klebsiella
40	10	23.3	7	2 Q47477	Q47477 escherichia
41	10	23.3	7	8 Q98866	Q98866 spinacia ol
42	10	23.3	7	10 Q9C5B3	Q9C5B3 arabidopsis
43	10	23.3	8	2 Q45889	Q45889 clostridium
44	10	23.3	8	3 P82858	P82858 puccinia re
45	10	23.3	8	4 Q9Y4X6	Q9Y4X6 homo sapien

ALIGNMENTS

RESULT 1

Q9PS69 PRELIMINARY; PRT; 8 AA.  
ID Q9PS69  
AC Q9PS69;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE LOW density lipoprotein receptor-related protein (Fragment).  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=92011685; PubMed=1918027;  
RA Stifani S., Barber D.L.; Aebersold R., Steyrer E., Shen X., Nimpf J.,  
RA Schneider W.J.;  
RT "The laying hen expresses two different low density lipoprotein  
RT receptor-related proteins.";  
RL J. Biol. Chem. 266:19079-19087(1991).  
FT NON\_TER 1 1  
FT NON\_TER 8 8  
SQ SEQUENCE 8 AA; 846 MW; C007272DD865BAAA CRC64;

Query Match 37.2%; Score 16; DB 13; Length 8;  
Best Local Similarity 75.0%; Pred. No. 8.3e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 GAML 9  
Db 4 GALL 7

RESULT 2

Q9BQT4 PRELIMINARY; PRT; 9 AA.  
ID Q9BQT4  
AC Q9BQT4;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE CLIP-associating protein CLASP2 (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RA Galjart N.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=21185938; PubMed=11290329;  
 RA Akhmanova A., Hoogenraad C.C., Drabek K., Stepanova T., Dortland B.,  
 RA Verkerk T., Vermeulen W., Burgering B.M., de Zeeuw C.I., Grosveld F.,  
 RA Galjart N.;  
 RT "CLASPs are CLIP-115 and -170 associating proteins involved in the  
 RT regional regulation of microtubule dynamics in motile fibroblasts.";  
 DR Cell 104:923-935(2001).  
 DR EMBL; AJ288058; CAC35157.1; -.  
 FT NON\_TER 9  
 SQ SEQUENCE 9 AA; 1001 MW; C3ECB33AAAA8769D CRC64;

Query Match 34.98; Score 15; DB 4; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 8.3e+05;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 MAQG 6  
 DB 1 MAMG 4

RESULT 3  
 Q9NYH5 PRELIMINARY; PRT; 9 AA.  
 AC Q9NYH5;  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-OCT-2000 (Tremblrel. 22, Last annotation update)  
 DE Hypothetical protein KIAA0020 (Fragment).  
 GN KIAA0020.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Akatsuka Y., Brickner A.G., Warren E.H., Engelhard V.H., Riddell S.R.;  
 RT "Determination of intronic sequences adjacent to exonic sequence using  
 RT PCR and a genomic DNA library constructed by TA cloning.";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF245442; AAF69838.1; -.  
 KW Hypothetical protein.  
 FT NON\_TER 1  
 FT NON\_TER 9  
 SQ SEQUENCE 9 AA; 1029 MW; 0A73DD68771A720 CRC64;

Query Match 32.6%; Score 14; DB 4; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 MAQ 5  
 DB 6 MAQ 8

RESULT 4  
 Q9TWD6 PRELIMINARY; PRT; 9 AA.  
 ID Q9TWD6  
 AC Q9TWD6;

DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)  
 DE LED-NPF-1-NEUROPEPTIDE F-related peptide.  
 OS Leptinotarsa decemlineata (Colorado potato beetle).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;  
 OC Phytophaga; Chrysomeloidea; Chrysomelidae; Chrysomelinae;  
 OC Chrysomelini; Leptinotarsa.  
 OX NCBI\_TaxID=7539;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=96245438; PubMed=8814784;  
 RA Spittaels K., Verhaert P., Shaw C., Johnston R.N., Devreese B.,  
 RA Van Beeumen J., De Loof A.;  
 RT "Insect neuropeptide F (NPF)-related peptides: isolation from Colorado  
 RT potato beetle (Leptinotarsa decemlineata) brain.";  
 RL Insect Biochem. Mol. Biol. 26:375-382(1996).  
 SQ SEQUENCE 9 AA; 1066 MW; 7E02340736C76864 CRC64;

Query Match 32.6%; Score 14; DB 5; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 8.3e+05;  
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 AOGAML 9  
 DB 1 ARGQL 6

RESULT 5  
 P83152 PRELIMINARY; PRT; 8 AA.  
 ID P83152;  
 AC P83152;  
 DT 01-MAR-2002 (Tremblrel. 20, Created)  
 DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)  
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
 DE Allophycocyanin beta chain (Fragment).  
 OS Anabaena sp. (Strain L31).  
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.  
 OX NCBI\_TaxID=29412;  
 RN [1]  
 RP SEQUENCE.  
 RA Apte S.K., Uhlemann E., Schmid R., Altendorf K.;  
 RL Submitted (OCT-2001) to the SWISS-PROT data bank.  
 CC -1- FUNCTION: LIGHT-HARVESTING PHOTOSYNTHETIC BILE PIGMENT-PROTEIN  
 CC FROM THE PHYCOBILIPROTEIN COMPLEX. ALLOPHYCOCYANIN HAS A MAXIMUM  
 CC ABSORPTION AT APPROXIMATELY 650 TO 653 NANOMETERS.  
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.  
 CC -1- SUBCELLULAR LOCATION: THYLAKOID MEMBRANE.  
 CC -1- PTM: CONTAINS ONE COVALENTLY LINKED BILIN CHROMOPHORE.  
 KW Phycobilisome; Electron transport; Photosynthesis; Bile pigment;  
 KW Thylakoid; Membrane.  
 FT NON\_TER 8  
 SQ SEQUENCE 8 AA; 788 MW; 87CDC1A05DDAB6DD CRC64;

Query Match 30.2%; Score 13; DB 2; Length 8;  
 Best Local Similarity 60.0%; Pred. No. 8.3e+05;  
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 AOGAM 8  
 DB 1 AQDAI 5

RESULT 6  
 Q8R4D8 PRELIMINARY; PRT; 8 AA.  
 ID Q8R4D8  
 AC Q8R4D8;  
 DT 01-JUN-2002 (Tremblrel. 21, Created)  
 DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE Inhibitory PAS domain protein (Fragment).  
 GN IPAS.

```
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Makino Y., Kanopka A., Berkenstam A., Tanaka H., Poellinger L.;
RT "Cloning of mouse inhibitory PAS domain protein gene.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF481145; AAL87226.1; -.
FT NON_TER 8
SQ SEQUENCE 8 AA; 887 MW; EBD416D728772DD6 CRC64;

Query Match 30.2%; Score 13; DB 11; Length 8;
Best Local Similarity 75.0%; Pred. No. 8.3e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 MAQG 6
DB 1 MALG 4

RESULT 7
Q83977 ID Q83977 PRELIMINARY; PRT; 8 AA.
AC Q83977;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Influenza A/udorn/72 (H3n2), nucleoprotein (Seg 5), 5' cdna
DE (Fragment).
OS Influenzavirus A.
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses.
OX NCBI_TaxID=197911;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83112211; PubMed=6296449;
RA Lin B.-C., Lai C.-J.;
RT "The influenza virus nucleoprotein synthesized from cloned dna in a
RT simian virus 40 vector is detected in the nucleus.";
RL J. Virol. 45:434-438(1983).
DR EMBL; J02170; AAA43468.1; -.
KW Nucleocapsid.
FT NON_TER 8
SQ SEQUENCE 8 AA; 878 MW; ED0321A866C5BDD6 CRC64;

Query Match 30.2%; Score 13; DB 12; Length 8;
Best Local Similarity 40.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 LMAQG 6
DB 1 MASOG 5

RESULT 8
Q83349 ID Q83349 PRELIMINARY; PRT; 8 AA.
AC Q83349;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical 0.8 kDa protein.
OS Murine coronavirus MHV (strain A59).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=11142;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A59;
RX MEDLINE=89299451; PubMed=2545027;
```

```
RA Pachuk C.J., Bredenbeek P.J., Zoltick P.W., Spaan W.J., Weiss S.R.;
RT "Molecular cloning of the gene encoding the putative polymerase of
RT mouse hepatitis coronavirus, strain A59.";
RL Virology 171:141-148(1989).
DR EMBL; M27198; AAA74010.1; -.
KW Hypothetical protein.
SQ SEQUENCE 8 AA; 787 MW; ECB732C7287DC766 CRC64;

Query Match 30.2%; Score 13; DB 12; Length 8;
Best Local Similarity 42.9%; Pred. No. 8.3e+05;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 MAQGAML 9
DB 1 MPAGLVL 7

RESULT 9
Q43960 ID Q43960 PRELIMINARY; PRT; 9 AA.
AC Q43960;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hydrogenase-related protein (Fragment).
GN HUPA.
OS Azotobacter chroococcum.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Azotobacter.
OX NCBI_TaxID=353;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MCd1;
RX MEDLINE=95055698; PubMed=7966281;
RA Du L., Tibelius K.H., Souza E.M., Garg R.P., Yates M.G.;
RT "Sequences, organization and analysis of the hup2MMOQRTV genes from
RT the Azotobacter chroococcum hydrogenase gene cluster.";
RL J. Mol. Biol. 243:549-557(1994).
DR EMBL; L25315; AAA64455.1; -.
FT NON_TER 9
SQ SEQUENCE 9 AA; 1004 MW; EF421DD045B69811 CRC64;

Query Match 30.2%; Score 13; DB 2; Length 9;
Best Local Similarity 50.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 MAQG 6
DB 6 IAEG 9

RESULT 10
P72149 ID P72149 PRELIMINARY; PRT; 9 AA.
AC P72149;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Putative glucokinase (Fragment).
GN GLK.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PA01;
RX MEDLINE=96427344; PubMed=8830708;
RA Sage A.E., Proctor W.D., Phibbs P.V.Jr.;
RT "A two-component response regulator, gltR, is required for glucose
RT transport activity in Pseudomonas aeruginosa PA01.";
RL J. Bacteriol. 178:6064-6066(1996).
DR EMBL; U50932; AAC44474.1; -;
```

```

KW Kinase.
FT NON_TER
SQ SEQUENCE 9 AA; 944 MW; C3071DDAA72DC6C6 CRC64;

Query Match
Best Local Similarity 30.2%; Score 13; DB 2; Length 9;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LMAOG 6
DB 5 LDAEG 9

RESULT 11
O95953
ID O95953 PRELIMINARY; PRT; 9 AA.
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-MAY-2002 (TReMBLrel. 20, Last annotation update)
DE Galactocerebrosidase (EC 3.2.1.46) (Fragment).
GN GALT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Lulli L., Torchiana E., Finocchiaro G.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U77631; AAD15626.1; -.
KW Glycosidase; Hydrolase.
FT NON_TER 9
SQ SEQUENCE 9 AA; 1069 MW; 374E2AADC2C699C8 CRC64;

Query Match
Best Local Similarity 30.2%; Score 13; DB 4; Length 9;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMA 4
DB 3 FMVA 6

RESULT 12
O9TLM7
ID O9TLM7 PRELIMINARY; PRT; 9 AA.
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE Ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit
DE (Fragment).
GN RBCS.
OS Laurencia viridis.
OG Chloroplast.
OC Eukaryota; Rhodophyta; Florideophyceae; Ceramiales; Rhodomelaceae;
OX Laurencia.
OX NCBI_TaxID=99904;
RN [1]
RP SEQUENCE FROM N.A.
RA Lewis S.M., Valdes F., Frias I., Hernandez-Gonzalez M.C.,
RA Gil-Rodriguez M.C., Gacesa P.;
RT "Analysis of rDNA and RUBISCO spacer sequences of the genera Laurencia
RT and Osmundea in the Canary Islands.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF081272; AAD53999.1; -.
KW Chloroplast.
FT NON_TER 9
SQ SEQUENCE 9 AA; 1040 MW; 7080D1A866C1A734 CRC64;

Query Match
Best Local Similarity 30.2%; Score 13; DB 8; Length 9;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLM 3
DB 6 FLL 8

Query Match
Best Local Similarity 27.9%; Score 12; DB 2; Length 8;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLM 3
DB 6 FLL 8

Best Local Similarity 50.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 MAQG 6
DB 3 LTQG 6

RESULT 13
O85599
ID O85599 PRELIMINARY; PRT; 9 AA.
AC O85599;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Moloney murine sarcoma virus (Strain HT1) mos/env 3' junction
DE (Fragment).
OS Moloney murine leukemia virus.
OC Viruses; Retroviruses; Retroviridae; Gammaretrovirus.
OX NCBI_TaxID=11801;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83164305; PubMed=6300424;
RA Donoghue D.J., Hunter T.;
RT "Recombination junctions of variants of Moloney murine sarcoma virus:
RT Generation and divergence of a mammalian transforming gene.";
RL J. Virol. 45:607-617(1983).
DR EMBL; K03107; AAA46493.1; -.
FT NON_TER 1
SQ SEQUENCE 9 AA; 932 MW; 410B2DD86409CDD3 CRC64;

Query Match
Best Local Similarity 30.2%; Score 13; DB 15; Length 9;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 QGAM 8
DB 5 RGAL 8

RESULT 14
O93SP2
ID O93SP2 PRELIMINARY; PRT; 8 AA.
AC O93SP2;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Quaternary ammonium compound-resistance protein QacEdelta1
DE (fragment).
GN QACDELTA1.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YMCT704; TRANSPOSON=class I integron;
RA Lee K., Chong Y., Yum J.H., Yong D., Livermore D.M.;
RT "VIM-2 metallo-beta-lactamase gene-containing integron in a
RT Pseudomonas aeruginosa clinical isolate.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY029772; AAK50441.1; -.
FT NON_TER 8
SQ SEQUENCE 8 AA; 920 MW; FE2729C732C87336 CRC64;

Query Match
Best Local Similarity 66.7%; Score 12; DB 2; Length 8;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLM 3
DB 6 FLL 8

```

```
RESULT 15
Q9UCN4
ID Q9UCN4 PRELIMINARY; PRT; 8 AA.
AC Q9UCN4;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Cell-surface heparin/HEPARANSULFATE-binding protein peptide 3
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=92291065; PubMed=1601862;
RA Raboudi N., Julian J., Rohde L.H., Carson D.D.;
RT "Identification of cell-surface heparin/heparan sulfate-binding
RT proteins of a human uterine epithelial cell line (RL95).";
RL J. Biol. Chem. 267:11930-11939(1992).
SQ SEQUENCE 8 AA; 689 MW; 80B8733DD3DD87D CRC64;

Query Match 27.9%; Score 12; DB 4; Length 8;
Best Local Similarity 50.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 AOGA 7
Db 1:1
5 AKGS 8
```

Search completed: October 7, 2003, 14:48:10  
Job time : 60.5 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:43:36 ; Search time 20 Seconds  
(without alignments)  
19.040 Million cell updates/sec

Title: US-09-807-512-24  
Perfect score: 43  
Sequence: 1 FLMAQGAML 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 77717

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	51.2	7	1	US-08-273-474-31
2	22	51.2	7	3	US-08-935-100-31
3	22	51.2	7	5	PCT-US94-07881-31
4	21	48.8	9	4	US-08-197-484-34
5	21	48.8	9	4	US-09-311-784A-175
6	21	48.8	9	5	PCT-US95-02121-34
7	20	46.5	8	4	US-08-884-235-10
8	20	46.5	9	3	US-08-159-339A-695
9	20	46.5	9	6	5190919-3
10	19	44.2	9	1	US-08-467-083-15
11	19	44.2	9	1	US-08-414-417B-15
12	19	44.2	9	2	US-08-486-348A-15
13	19	44.2	9	2	US-08-468-545B-15
14	19	44.2	9	2	US-08-146-028-308
15	19	44.2	9	2	US-08-146-028-309
16	19	44.2	9	2	US-08-146-028-314
17	19	44.2	9	2	US-08-146-028-315
18	19	44.2	9	2	US-08-146-028-316
19	19	44.2	9	2	US-08-146-028-317
20	19	44.2	9	2	US-08-146-028-318
21	19	44.2	9	2	US-08-146-028-322
22	19	44.2	9	2	US-08-146-028-323
23	19	44.2	9	2	US-08-146-028-324
24	19	44.2	9	2	US-08-986-234-95
25	19	44.2	9	3	US-08-466-680B-15
26	19	44.2	9	3	US-08-723-425A-308
27	19	44.2	9	3	US-08-723-425A-309

28	19	44.2	9	3	US-08-723-425A-314	Sequence 314, App
29	19	44.2	9	3	US-08-723-425A-315	Sequence 315, App
30	19	44.2	9	3	US-08-723-425A-316	Sequence 316, App
31	19	44.2	9	3	US-08-723-425A-317	Sequence 317, App
32	19	44.2	9	3	US-08-723-425A-318	Sequence 318, App
33	19	44.2	9	3	US-08-723-425A-322	Sequence 322, App
34	19	44.2	9	3	US-08-723-425A-323	Sequence 323, App
35	19	44.2	9	3	US-08-723-425A-324	Sequence 324, App
36	19	44.2	9	3	US-09-112-206-308	Sequence 308, App
37	19	44.2	9	3	US-09-112-206-309	Sequence 309, App
38	19	44.2	9	3	US-09-112-206-314	Sequence 314, App
39	19	44.2	9	3	US-09-112-206-315	Sequence 315, App
40	19	44.2	9	3	US-09-112-206-316	Sequence 316, App
41	19	44.2	9	3	US-09-112-206-317	Sequence 317, App
42	19	44.2	9	3	US-09-112-206-318	Sequence 318, App
43	19	44.2	9	3	US-09-112-206-322	Sequence 322, App
44	19	44.2	9	3	US-09-112-206-323	Sequence 323, App
45	19	44.2	9	3	US-09-112-206-324	Sequence 324, App

ALIGNMENTS

RESULT 1  
US-08-273-474-31  
; Sequence 31, Application US/08273474 ;  
; Patent No. 5691134  
; GENERAL INFORMATION:  
; APPLICANT: Kilpatrick, David R.  
; TITLE OF INVENTION: POLIOVIRUS SPECIFIC PRIMERS AND METHODS  
; TITLE OF INVENTION: OF DETECTION UTILIZING THE SAME  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NEEDLE & ROSENBERG, P.C.  
; STREET: Suite 1200, The Candler-Building, 127  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA  
; ZIP: 30303-1811  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/273,474  
; FILING DATE: 13-JUL-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Perryman, David G.  
; REGISTRATION NUMBER: 33,438  
; REFERENCE/DOCKET NUMBER: 1414.617  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (404) 688-9880  
; TELEFAX: (404) 688-9870  
; INFORMATION FOR SEQ ID NO: 31:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-273-474-31

Query Match 51.2%; Score 22; DB 1; Length 7;  
Best Local Similarity 66.7%; Pred. No. 2.5e+05;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 3 MAQAM 8  
: : : : :  
DB 2 VAQAL 7

RESULT 2  
US-08-935-100-31  
; Sequence 31, Application US/08935100  
; Patent No. 6143494  
; GENERAL INFORMATION:  
; APPLICANT: Kilpatrick, David R.  
; TITLE OF INVENTION: POLIOVIRUS SPECIFIC PRIMERS AND METHODS  
; TITLE OF INVENTION: OF DETECTION UTILIZING THE SAME  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NEEDLE & ROSENBERG, P.C.  
; STREET: Suite 1200, The Candler Building, 127  
; STREET: Peachtree Street, NE  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA  
; ZIP: 30303-1811  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/08/935,100  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA: US/08/273,474  
; FILING DATE: 13-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Perryman, David G.  
; REGISTRATION NUMBER: 33,438  
; REFERENCE/DOCKET NUMBER: 1414.617  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (404) 688-0770  
; TELEFAX: (404) 688-9880  
; INFORMATION FOR SEQ ID NO: 31:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-935-100-31

Query Match 51.2%; Score 22; DB 3; Length 7;  
Best Local Similarity 66.7%; Pred. No. 2.5e+05;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 MAQAM 8  
; : : : :  
Db 2 VAQAL 7

RESULT 3  
PCT-US94-07881-31  
; Sequence 31, Application PC/TUS9407881  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of  
; APPLICANT: Department of Health and Human Services  
; APPLICANT: National Institutes of Health  
; APPLICANT: Office of Technology Transfer  
; APPLICANT: Bethesda, Maryland 20892  
; TITLE OF INVENTION: POLIOVIRUS SPECIFIC PRIMERS AND  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NEEDLE & ROSENBERG, P.C.  
; STREET: Suite 1200, The Candler Building,  
; STREET: 127 Peachtree Street NE  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA

America, as re

; ZIP: 30303-1811  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/07881  
; FILING DATE: 13 July 1994  
; PRIOR APPLICATION DATA: US 08/092,110  
; APPLICATION NUMBER: US 08/092,110  
; FILING DATE: 13 July 1993  
; INFORMATION FOR SEQ ID NO: 31:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
PCT-US94-07881-31

Query Match 51.2%; Score 22; DB 5; Length 7;  
Best Local Similarity 66.7%; Pred. No. 2.5e+05;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 MAQAM 8  
; : : : :  
Db 2 VAQAL 7

RESULT 4  
US-08-197-484-34  
; Sequence 34, Application US/08197484  
; Patent No. 6419931  
; GENERAL INFORMATION:  
; APPLICANT: VITIELLO, Maria A.  
; APPLICANT: CHESTNUT, Robert W.  
; APPLICANT: SETTE, Alessandro D.  
; APPLICANT: CELIS, Esteban  
; APPLICANT: GRAY, Howard  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING  
; TITLE OF INVENTION: CTL IMMUNITY  
; NUMBER OF SEQUENCES: 153  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourie and Crew  
; STREET: Stewart Street Tower, One Market Plaza  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: US  
; ZIP: 94105-1493  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/197,484  
; FILING DATE: 16-FEB-1994  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA: US 07/935,811  
; APPLICATION NUMBER: US 07/935,811  
; FILING DATE: 26-AUG-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/874,491  
; FILING DATE: 27-APR-1992  
; PRIOR APPLICATION DATA: US 07/827,682  
; APPLICATION NUMBER: US 07/827,682  
; FILING DATE: 29-JAN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/749,568  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Parmelee, Steven W.



REGISTRATION NUMBER: 31,990  
REFERENCE/DOCKET NUMBER: 14137-26-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 467-9600  
TELEFAX: (206) 623-6793  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-197-484-34

Query Match 48.8%; Score 21; DB 4; Length 9;  
Best Local Similarity 80.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAQ 5  
||:|  
Db 1 FLFAQ 5

RESULT 5  
US-09-311-784A-175  
Sequence 175, Application US/09311784A  
Patent No. 6534482  
GENERAL INFORMATION:  
APPLICANT: Fikes, John D.  
APPLICANT: Hermanson, Gary G.  
APPLICANT: Sette, Alessandro  
APPLICANT: Ishioka, Glenn Y.  
APPLICANT: Livingston, Brian  
APPLICANT: Chesnut, Robert W.  
APPLICANT: Epimmune Inc.  
TITLE OF INVENTION: Expression Vectors for Stimulating an  
FILE OF INVENTION: Immune Response and Methods of Using the Same  
FILE REFERENCE: 39963-20022.01  
CURRENT APPLICATION NUMBER: US/09/311,784A  
CURRENT FILING DATE: 1999-05-13  
PRIOR APPLICATION NUMBER: US 60/085,751  
PRIOR FILING DATE: 1998-05-15  
NUMBER OF SEQ ID NOS: 463  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 175  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: HBV pol 503 (peptide 1069.071)  
US-09-311-784A-175

Query Match 48.8%; Score 21; DB 4; Length 9;  
Best Local Similarity 80.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAQ 5  
||:|  
Db 1 FLFAQ 5

RESULT 6  
PCT-US95-02121-34  
Sequence 34, Application PC/TUS9502121  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING  
CTL IMMUNITY  
NUMBER OF SEQUENCES: 153  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/02121  
FILING DATE: 16-FEB-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/197,484  
FILING DATE: 16-FEB-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/935,811  
FILING DATE: 26-AUG-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/874,491  
FILING DATE: 27-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/827,682  
FILING DATE: 29-JAN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/749,568  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Parmelee, Steven W.  
REGISTRATION NUMBER: 31,990  
REFERENCE/DOCKET NUMBER: 14137-26-4PC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 467-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
PCT-US95-02121-34

Query Match 48.8%; Score 21; DB 5; Length 9;  
Best Local Similarity 80.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAQ 5  
||:|  
Db 1 FLFAQ 5

RESULT 7  
US-08-884-235-10  
Sequence 10, Application US/08884235  
Patent No. 6329573  
GENERAL INFORMATION:  
APPLICANT: Lightfoot, David A.  
APPLICANT: Long, Lynn M.  
TITLE OF INVENTION: PLANTS CONTAINING THE gdhA GENE AND  
METHODS OF USE THEREOF  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.  
STREET: 1100 New York Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: United States of America  
ZIP: 20005-3918  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: MS Word  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/884,235  
FILING DATE: 27-JUN-1997  
CLASSIFICATION: 800  
INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:  
LENGTH: 8 residues  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: protein  
US-08-884-235-10

Query Match 46.5%; Score 20; DB 4; Length 8;  
Best Local Similarity 42.9%; Pred. No. 2.5e+05;  
Matches 3; Conservative 3; Mismatches 1; Indels 0;

Qy 2 LMAQAM 8  
Db 2 LMAQVI 8

## RESULT 8

US-08-159-339A-695  
; Sequence 695, Application US/08159339A  
; Patent No. 6037135  
; GENERAL INFORMATION:  
; APPLICANT: Kubo, Ralph T.  
; APPLICANT: Grey, Howard M.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Cellis, Esteban  
; TITLE OF INVENTION: HLA Binding peptides and Their  
; NUMBER OF SEQUENCES: 1254  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/159,339A  
; FILING DATE: 29-NOV-1993  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/926,666  
; FILING DATE: 07-AUG-1992  
; APPLICATION NUMBER: US 08/027,746  
; FILING DATE: 05-MAR-1993  
; APPLICATION NUMBER: US 08/103,396  
; FILING DATE: 06-AUG-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weber, Ellen Lauver  
; REGISTRATION NUMBER: 32,762  
; REFERENCE/DOCKET NUMBER: 018623-005030US  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 695:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-159-339A-695

Query Match 46.5%; Score 20; DB 3; Length 9;  
Best Local Similarity 44.4%; Pred. No. 2.5e+05;  
Matches 4; Conservative 1; Mismatches 4; Indels 0;

Qy 1 FLMAQAML 9  
Db 1 FFFLYGALL 9

## RESULT 9

5190919-3  
; Patent No. 5190919  
; APPLICANT: FALLADINO, LINDA O.; SILBERKLANG, MELVIN; TUNG,  
; JWU-SHENG; LAW, SIMON W.; MARK, GEORGE E.  
; TITLE OF INVENTION: ANTISTAIN DERIVED ANTICOAGULANT PROTEIN  
; NUMBER OF SEQUENCES: 40  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/512,695  
; FILING DATE: 23-APR-1990  
; SEQ ID NO: 3  
; LENGTH: 9  
5190919-3

Query Match 46.5%; Score 20; DB 6; Length 9;  
Best Local Similarity 50.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 LMAQGAML 9  
Db 1 LLYNGAQL 8

## RESULT 10

US-08-467-083-15  
; Sequence 15, Application US/08467083  
; Patent No. 5726023  
; GENERAL INFORMATION:  
; APPLICANT: Cheever, Martin A.  
; APPLICANT: Disis, Mary L.  
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/NEU PROTEIN  
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
; TITLE OF INVENTION: HER-2/NEU ONCOGENE IS ASSOCIATED  
; NUMBER OF SEQUENCES: 68  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed and Berry  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: US  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/467,083  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/414,417  
; FILING DATE: 06-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharkey, Richard G.  
; REGISTRATION NUMBER: 32,629  
; REFERENCE/DOCKET NUMBER: 920010.448C2  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; TELEX: 3723836 SEEDANBERRY  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-467-083-15

```
Query Match      44.2%; Score 19; DB 1; Length 9;
Best Local Similarity 50.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      2 LMAOGAML 9
      || | | : |
Db      2 LMPYGCLL 9

RESULT 11
US-08-414-417B-15
; Sequence 15, Application US/08414417B
; Patent No. 5801005
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disis, Mary L.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/414,417B
; FILING DATE: 31-MAR-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-414-417B-15

Query Match      44.2%; Score 19; DB 1; Length 9;
Best Local Similarity 50.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      2 LMAOGAML 9
      || | | : |
Db      2 LMPYGCLL 9

RESULT 12
US-08-486-348A-15
; Sequence 15, Application US/08486348A
; Patent No. 5846538
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disis, Mary L.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
```

```
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,348A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-486-348A-15

Query Match      44.2%; Score 19; DB 2; Length 9;
Best Local Similarity 50.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      2 LMAOGAML 9
      || | | : |
Db      2 LMPYGCLL 9

RESULT 13
US-08-468-545B-15
; Sequence 15, Application US/08468545B
; Patent No. 5876712
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disis, Mary L.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,545B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
```

```
;
; TYPE: amino acid
; TOPOLOGY: linear
US-08-468-545B-15

Query Match          44.2%; Score 19; DB 2; Length 9;
Best Local Similarity 50.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 LMAQGAML 9
Db 2 LMPYGCLL 9

RESULT 14
US-08-146-028-308
; Sequence 308, Application US/08146028
; Patent No. 5891640
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; INFORMATION FOR SEQ ID NO: 308:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-146-028-308

Query Match          44.2%; Score 19; DB 2; Length 9;
Best Local Similarity 80.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 QGAML 9
Db 5 QGMML 9

RESULT 15
US-08-146-028-309
; Sequence 309, Application US/08146028
; Patent No. 5891640
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; INFORMATION FOR SEQ ID NO: 309:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
```

```
;
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-146-028-309

Query Match          44.2%; Score 19; DB 2; Length 9;
Best Local Similarity 80.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 QGAML 9
Db 4 QGMML 8

Search completed: October 7, 2003, 14:49:51
Job time : 20 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:45:41 ; Search time 21 Seconds  
(without alignments)  
67.806 Million cell updates/sec

Title: US-09-807-512-24  
Perfect score: 43  
Sequence: 1 FLMAQGAML 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues  
Total number of hits satisfying chosen parameters: 59020

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PublishedApplications\_AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	24	55.8	9	12	US-09-932-165-266
2	21	48.8	9	15	US-10-128-711-34
3	20	46.5	9	11	US-09-809-638-454
4	20	46.5	9	12	US-09-932-165-75
5	20	46.5	9	12	US-10-210-148-9
6	20	46.5	9	12	US-10-210-148-14
7	19	44.2	9	8	US-08-812-393A-55
8	19	44.2	9	9	US-09-810-936-310
9	19	44.2	9	10	US-09-429-755-310
10	19	44.2	9	10	US-09-924-400-310
11	19	44.2	9	12	US-10-283-618-1
12	19	44.2	9	12	US-10-283-618-3
13	19	44.2	9	12	US-10-353-929-28
14	19	44.2	9	15	US-10-001-546-9
15	19	44.2	9	15	US-10-086-082-6

16	19	44.2	9	15	US-10-212-679-310	Sequence 310, App
17	18	41.9	8	9	US-09-739-907-120	Sequence 120, App
18	18	41.9	8	9	US-09-801-784-24	Sequence 24, Appl
19	18	41.9	8	9	US-09-884-681-26	Sequence 26, Appl
20	18	41.9	8	12	US-10-293-580-26	Sequence 26, Appl
21	18	41.9	9	9	US-09-862-260A-21	Sequence 21, Appl
22	18	41.9	9	9	US-09-796-744-10	Sequence 10, Appl
23	18	41.9	9	10	US-09-894-018-290	Sequence 290, App
24	18	41.9	9	10	US-09-920-174-11	Sequence 11, Appl
25	18	41.9	9	10	US-09-909-460-94	Sequence 94, Appl
26	18	41.9	9	12	US-10-359-431-5	Sequence 5, Appli
27	18	41.9	9	12	US-10-359-431-19	Sequence 19, Appl
28	18	41.9	9	12	US-09-920-195A-11	Sequence 11, Appl
29	18	41.9	9	12	US-10-231-452-7	Sequence 7, Appli
30	18	41.9	9	15	US-10-128-711-35	Sequence 35, Appl
31	18	41.9	9	15	US-10-133-210-105	Sequence 105, App
32	18	41.9	9	15	US-10-146-505-91	Sequence 91, Appl
33	17	39.5	8	9	US-09-796-264-6	Sequence 6, Appli
34	17	39.5	8	9	US-09-796-264-9	Sequence 9, Appli
35	17	39.5	8	10	US-09-845-226-6	Sequence 6, Appli
36	17	39.5	8	10	US-09-845-226-9	Sequence 9, Appli
37	17	39.5	8	10	US-09-998-831-41	Sequence 41, Appl
38	17	39.5	8	10	US-09-795-903A-6	Sequence 6, Appli
39	17	39.5	8	10	US-09-795-903A-9	Sequence 9, Appli
40	17	39.5	8	12	US-10-079-167-60	Sequence 60, Appl
41	17	39.5	8	12	US-10-373-561-41	Sequence 41, Appl
42	17	39.5	8	12	US-10-280-066-204	Sequence 204, App
43	17	39.5	8	15	US-10-032-818-9	Sequence 9, Appli
44	17	39.5	8	15	US-10-032-818-11	Sequence 11, Appl
45	17	39.5	8	15	US-10-168-843A-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1  
US-09-932-165-266  
; Sequence 266, Application US/09932165  
; Publication No. US20030134784A1  
; GENERAL INFORMATION:  
; APPLICANT: RAITANO, ARTHUR  
; APPLICANT: CHALLITA-EID, PIA M.  
; APPLICANT: FARIS, MARY  
; APPLICANT: SAFFRAN, DOUGLAS  
; APPLICANT: AFAR, DANIEL  
; APPLICANT: LEVIN, ELANA  
; APPLICANT: HUBERT, RENE  
; APPLICANT: GE, WANGMAO  
; APPLICANT: JAKOBOVITS, AYA  
; TITLE OF INVENTION: NUCLEIC ACIDS AND CORRESPONDING PROTEINS ENTITLED  
; TITLE OF INVENTION: 83P2H3 AND Cat-F2E11 USEFUL IN TREATMENT AND  
; TITLE OF INVENTION: DETECTION OF CANCER  
; FILE REFERENCE: 51158-20014.00  
; CURRENT APPLICATION NUMBER: US/09/932,165  
; CURRENT FILING DATE: 2001-08-17  
; PRIOR APPLICATION NUMBER: 60/226,329  
; PRIOR FILING DATE: 2000-08-17  
; NUMBER OF SEQ ID NOS: 1508  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 266  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Peptide motif  
US-09-932-165-266

Query Match 55.8%; Score 24; DB 12; Length 9;  
Best Local Similarity 83.3%; Pred. No. 5, 3e+05;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 2 LMAQGA 7  
|:||||

Db 2 LVAQGA 7

## RESULT 2

US-10-128-711-34  
; Sequence 34, Application US/10128711  
; Publication No. US2003009634A1

## GENERAL INFORMATION:

APPLICANT: VITIELLO, Maria A.  
CHESTNUT, Robert W.  
SETTE, Alessandro D.  
CELIS, Esteban  
GRAY, Howard

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING

CTL IMMUNITY

NUMBER OF SEQUENCES: 153

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend Kourie and Crew

STREET: Stewart Street Tower, One Market Plaza

CITY: San Francisco

STATE: California

COUNTRY: US

ZIP: 94105-1493

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/128,711

FILING DATE: 22-Apr-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/197,484

FILING DATE: 16-FEB-1994

APPLICATION NUMBER: US 07/935,811

FILING DATE: 26-AUG-1992

APPLICATION NUMBER: US 07/874,491

FILING DATE: 27-APR-1992

APPLICATION NUMBER: US 07/827,682

FILING DATE: 29-JAN-1992

APPLICATION NUMBER: US 07/749,568

FILING DATE: 26-AUG-1991

ATTORNEY/AGENT INFORMATION:

NAME: Parmelee, Steven W.

REGISTRATION NUMBER: 31,990

REFERENCE/DOCKET NUMBER: 14137-26-4

TELEPHONE: (206) 467-9600

TELEFAX: (206) 623-6793

INFORMATION FOR SEQ ID NO: 34:

SEQUENCE CHARACTERISTICS:

LENGTH: 9 amino acids

TYPE: amino acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 34:

US-10-128-711-34

Query Match 48.8%; Score 21; DB 15; Length 9;  
Best Local Similarity 80.0%; Pred. No. 5.3e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAQ 5

Db 1 FLAQ 5

## RESULT 3

US-09-809-638-454

; Sequence 454, Application US/09809638

; Publication No. US20030059895A1

## GENERAL INFORMATION:

APPLICANT: Mary Paris  
APPLICANT: Pia M. Challita-Eid  
APPLICANT: Steve Chappell Mitchell

APPLICANT: Daniel E.H. Afar

APPLICANT: Arthur B. Raitano

APPLICANT: Aya Jakobovits

TITLE OF INVENTION: 125P5C8: A TISSUE SPECIFIC PROTEIN

TITLE OF INVENTION: HIGHLY EXPRESSED IN VARIOUS CANCERS

FILE REFERENCE: 129.350S01

CURRENT APPLICATION NUMBER: US/09/809,638

CURRENT FILING DATE: 2001-03-14

NUMBER OF SEQ ID NOS: 746

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 454

LENGTH: 9

TYPE: PRT

ORGANISM: Homo sapiens

US-09-809-638-454

Query Match 46.5%; Score 20; DB 11; Length 9;

Best Local Similarity 57.1%; Pred. No. 5.3e+05;

Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 MAQGA 9

Db 3 LASGLML 9

## RESULT 4

US-09-932-165-75

; Sequence 75, Application US/09932165

; Publication No. US20030134784A1

GENERAL INFORMATION:

APPLICANT: RAITANO, ARTHUR

APPLICANT: CHALLITA-EID, PIA M.

APPLICANT: FARIS, MARY

APPLICANT: SAFFRAN, DOUGLAS

APPLICANT: AFAR, DANIEL

APPLICANT: LEVIN, ELANA

APPLICANT: HUBERT, RENE

APPLICANT: GE, WANGMAO

APPLICANT: JAKOBOVITS, AYA

TITLE OF INVENTION: NUCLEIC ACIDS AND CORRESPONDING PROTEINS ENTITLED

TITLE OF INVENTION: 83P2H3 AND CatF2E11 USEFUL IN TREATMENT AND

TITLE OF INVENTION: DETECTION OF CANCER

FILE REFERENCE: 51158-20014.00

CURRENT APPLICATION NUMBER: US/09/932,165

CURRENT FILING DATE: 2001-08-17

PRIOR APPLICATION NUMBER: 60/226,329

PRIOR FILING DATE: 2000-08-17

NUMBER OF SEQ ID NOS: 1508

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 75

LENGTH: 9

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Peptide motif

US-09-932-165-75

Query Match 46.5%; Score 20; DB 12; Length 9;

Best Local Similarity 80.0%; Pred. No. 5.3e+05;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LMAQG 6

Db 5 LVAQG 9

## RESULT 5

US-10-210-148-9

; Sequence 9, Application US/10210148

```
; Publication No. US20030171280A1
; GENERAL INFORMATION:
; APPLICANT: Soderstrom, Karl Petter
; TITLE OF INVENTION: Compositions And Methods For Modulation Of Immune Response
; FILE REFERENCE: TROM0002
; CURRENT APPLICATION NUMBER: US/10/210,148
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: PCT/US02/24311
; PRIOR FILING DATE: 2002-07-31
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 9
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-210-148-9

Query Match 46.5%; Score 20; DB 12; Length 9;
Best Local Similarity 62.5%; Pred. No. 5.3e+05;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 LMAQGAML 9
   || || |
Db 1 LMLQGVDL 8

RESULT 6
US-10-210-148-14
; Sequence 14, Application US/10210148
; Publication No. US20030171280A1
; GENERAL INFORMATION:
; APPLICANT: Soderstrom, Karl Petter
; TITLE OF INVENTION: Compositions And Methods For Modulation Of Immune Response
; FILE REFERENCE: TROM0002
; CURRENT APPLICATION NUMBER: US/10/210,148
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: PCT/US02/24311
; PRIOR FILING DATE: 2002-07-31
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 14
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-210-148-14

Query Match 46.5%; Score 20; DB 12; Length 9;
Best Local Similarity 62.5%; Pred. No. 5.3e+05;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 LMAQGAML 9
   || || |
Db 1 LMLQGVDL 8

RESULT 7
US-08-812-393A-55
; Sequence 55, Application US/08812393A
; Publication No. US20010007152A1
; GENERAL INFORMATION:
; APPLICANT: SHERMAN, Linda A.
; APPLICANT: LUSTGARTEN, Joseph
; TITLE OF INVENTION: RECOMBINANT CONSTRUCTS ENCODING
; TITLE OF INVENTION: T CELL RECEPTORS SPECIFIC FOR HUMAN HLA-RESTRICTED TUMOR
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FORSTER
; STREET: 2000 Pennsylvania Avenue, NW, suite 5500
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/812,393A
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 31333-20001.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-887-1500
; TELEFAX: 202-822-0168
; TELEX:
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-812-393A-55

Query Match 44.2%; Score 19; DB 8; Length 9;
Best Local Similarity 50.0%; Pred. No. 5.3e+05;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 LMAQGAML 9
   || | | |
Db 2 LMPYGCLL 9

RESULT 8
US-09-810-936-310
; Sequence 310, Application US/09810936
; Patent No. US20020068285A1
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; APPLICANT: Reed, Steven G.
; APPLICANT: Smith, John M.
; APPLICANT: Misher, Linda E.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Day, Craig H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.419C11
; CURRENT APPLICATION NUMBER: US/09/810,936
; CURRENT FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 310
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Made in the lab
US-09-810-936-310

Query Match 44.2%; Score 19; DB 9; Length 9;
Best Local Similarity 50.0%; Pred. No. 5.3e+05;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 LMAQGAML 9
   || | | |
```

Db 2 LMAKALLL 9

## RESULT 9

US-09-429-755-310

; Sequence 310, Application US/09429755A  
; Patent No. US20020111467A1  
; GENERAL INFORMATION:  
; APPLICANT: Frudakis, Tony N.  
; APPLICANT: Smith, John M.  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Misher, Lynda  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Dillon, Davin C.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TREATMENT AND DIAGNOSIS OF BREAST CANCER  
; FILE REFERENCE: 210121.419C6  
; CURRENT APPLICATION NUMBER: US/09/429,755A  
; CURRENT FILING DATE: 1999-10-28  
; NUMBER OF SEQ ID NOS: 315  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 310  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Made in the lab

US-09-429-755-310

Query Match 44.2%; Score 19; DB 10; Length 9;  
Best Local Similarity 50.0%; Pred. No. 5.3e+05;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LMAQGAML 9

Db 2 LMAKALLL 9

## RESULT 10

US-09-924-400-310

; Sequence 310, Application US/09924400  
; Patent No. US20020165371A1  
; GENERAL INFORMATION:  
; APPLICANT: Frudakis, Tony N.  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Smith, John M.  
; APPLICANT: Misher, Lynda E.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A. W.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Deng, Ta  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; AND DIAGNOSIS OF BREAST CANCER  
; FILE REFERENCE: 210121.419C12  
; CURRENT APPLICATION NUMBER: US/09/924,400  
; CURRENT FILING DATE: 2001-08-07  
; NUMBER OF SEQ ID NOS: 340  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 310  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Made in the lab

US-09-924-400-310

Query Match 44.2%; Score 19; DB 10; Length 9;  
Best Local Similarity 50.0%; Pred. No. 5.3e+05;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LMAQGAML 9

Db 2 LMAKALLL 9

## RESULT 11

US-10-283-618-1  
; Sequence 1, Application US/10283618  
; Publication No. US20030165517A1  
; GENERAL INFORMATION:  
; APPLICANT: Nicolette, Charles A.  
; APPLICANT: Walker, Bruce  
; TITLE OF INVENTION: THERAPEUTIC ANTI-HIV (vpr) COMPOUNDS  
; FILE REFERENCE: GZ 2111.00  
; CURRENT APPLICATION NUMBER: US/10/283,618  
; CURRENT FILING DATE: 2002-10-29  
; PRIOR APPLICATION NUMBER: 60/345,957  
; PRIOR FILING DATE: 2001-10-29  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Human Immunodeficiency Virus  
; OTHER INFORMATION: Made in the lab

US-10-283-618-1

Query Match 44.2%; Score 19; DB 12; Length 9;  
Best Local Similarity 66.7%; Pred. No. 5.3e+05;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLMAQG 6

Db 1 FLYEQG 6

## RESULT 12

US-10-283-618-3

; Sequence 3, Application US/10283618  
; Publication No. US20030165517A1  
; GENERAL INFORMATION:  
; APPLICANT: Nicolette, Charles A.  
; APPLICANT: Walker, Bruce  
; TITLE OF INVENTION: THERAPEUTIC ANTI-HIV (vpr) COMPOUNDS  
; FILE REFERENCE: GZ 2111.00  
; CURRENT APPLICATION NUMBER: US/10/283,618  
; CURRENT FILING DATE: 2002-10-29  
; PRIOR APPLICATION NUMBER: 60/345,957  
; PRIOR FILING DATE: 2001-10-29  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Human Immunodeficiency Virus  
; OTHER INFORMATION: Made in the lab

US-10-283-618-3

Query Match 44.2%; Score 19; DB 12; Length 9;  
Best Local Similarity 66.7%; Pred. No. 5.3e+05;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLMAQG 6

Db 1 FLYEQG 6

## RESULT 13

US-10-353-929-28

; Sequence 28, Application US/10353929  
; Publication No. US20030175288A1  
; GENERAL INFORMATION:  
; APPLICANT: ITOH, Kyogo  
; TITLE OF INVENTION: Tumor antigen



```
; FILE REFERENCE: GP01-1024
; CURRENT APPLICATION NUMBER: US/10/353,929
; CURRENT FILING DATE: 2003-01-30
; PRIOR APPLICATION NUMBER: JP P2000-231814
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Designed peptide having an ability to activate HLA-A2 restricted
; OTHER INFORMATION: Cytotoxic T lymphocytes
US-10-353-929-28
```

```
Query Match 44.2%; Score 19; DB 12; Length 9;
Best Local Similarity 57.1%; Pred. No. 5.3e+05;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 2 LMAQGAM 8
||| |
Db 2 LMLQNAL 8
```

```
RESULT 14
US-10-001-546-9
; Sequence 9, Application US/10001546
; Publication No. US20030027766A1
; GENERAL INFORMATION:
; APPLICANT: IOANNIDES, CONSTANTIN G.
; APPLICANT: FISK, BRYAN A.
; APPLICANT: IOANNIDES, MARIA G.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STIMULATING
; TITLE OF INVENTION: T-LYMPHOCYTES
; FILE REFERENCE: UTSC:390USC2
; CURRENT APPLICATION NUMBER: US/10/001,546
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 08/403,459
; PRIOR FILING DATE: 1995-03-14
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-10-001-546-9
```

```
Query Match 44.2%; Score 19; DB 15; Length 9;
Best Local Similarity 50.0%; Pred. No. 5.3e+05;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 2 LMAQGAML 9
||| |
Db 2 LMPYGCLL 9
```

```
RESULT 15
US-10-086-082-6
; Sequence 6, Application US/10086082
; Publication No. US20030087402A1
; GENERAL INFORMATION:
; APPLICANT: Brieden, Walter
; APPLICANT: Naughton, Andrew
; APPLICANT: Robins, Karen
; APPLICANT: Shaw, Nicholas
; APPLICANT: Tinschert, Andreas
; APPLICANT: Zimmermann, Thomas
; TITLE OF INVENTION: METHOD OF PREPARING (S)-OR (R)
; TITLE OF INVENTION: -3,3-TRIFLUORO-2-HYDROXY-2-METHYLPROPIONIC ACID
```

```
; FILE REFERENCE: 32213
; CURRENT APPLICATION NUMBER: US/10/086,082
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US/09/214,679
; PRIOR FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Klebsiella oxytoca
US-10-086-082-6

Query Match 44.2%; Score 19; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 GAML 9
||| |
Db 4 GAML 7

Search completed: October 7, 2003, 14:50:40
Job time : 21 secs
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:46:12 ; Search time 277.5 Seconds

(without alignments)  
29.511 Million cell updates/sec

Title: US-09-807-512-24

Perfect score: 43

Sequence: 1 FLMAQGAML 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues

Total number of hits satisfying chosen parameters: 435500

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main :

1: /cgn2\_6/ptodata/1/paa/PCTUS\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US080\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep.\*  
8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep.\*  
9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep.\*  
10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep.\*  
11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep.\*  
12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep.\*  
13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep.\*  
14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep.\*  
15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep.\*  
16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep.\*  
17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep.\*  
18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep.\*  
19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep.\*  
20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep.\*  
21: /cgn2\_6/ptodata/1/paa/US097A\_COMB.pep.\*  
22: /cgn2\_6/ptodata/1/paa/US097B\_COMB.pep.\*  
23: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep.\*  
24: /cgn2\_6/ptodata/1/paa/US099A\_COMB.pep.\*  
25: /cgn2\_6/ptodata/1/paa/US099B\_COMB.pep.\*  
26: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep.\*  
27: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep.\*  
28: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep.\*  
29: /cgn2\_6/ptodata/1/paa/US103\_COMB.pep.\*  
30: /cgn2\_6/ptodata/1/paa/US104\_COMB.pep.\*  
31: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep.\*  
32: /cgn2\_6/ptodata/1/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query Score	Match Length	ID	Description
1	43	100.0	9 23 US-09-807-512-24	Sequence 24, Appl

2	24	55.8	8 20	US-09-641-377-1005	Sequence 1005, Ap
3	24	55.8	8 20	US-09-641-528A-3437	Sequence 3437, Ap
4	24	55.8	8 20	US-09-641-528A-18325	Sequence 18325, A
5	24	55.8	8 20	US-09-641-528A-25548	Sequence 25548, A
6	24	55.8	8 20	US-09-641-528A-33046	Sequence 33046, A
7	24	55.8	9 20	US-09-641-528A-18326	Sequence 18326, A
8	24	55.8	9 20	US-09-641-528A-25549	Sequence 25549, A
9	24	55.8	9 20	US-09-641-528A-33047	Sequence 33047, A
10	24	55.8	9 20	US-09-641-528A-40821	Sequence 40821, A
11	24	55.8	9 24	US-09-932-165-266	Sequence 266, App
12	24	55.8	9 27	US-10-182-252A-522	Sequence 522, App
13	23	53.5	8 26	US-10-031-874A-124	Sequence 124, App
14	23	53.5	8 26	US-10-031-874A-125	Sequence 125, App
15	23	53.5	8 26	US-10-031-874A-126	Sequence 126, App
16	21	48.8	8 1	PCT-US00-17842-472	Sequence 472, App
17	21	48.8	8 1	PCT-US00-24802-1169	Sequence 1169, Ap
18	21	48.8	8 1	PCT-US00-24802-1661	Sequence 1661, Ap
19	21	48.8	8 1	PCT-US00-24802-3341	Sequence 3341, Ap
20	21	48.8	8 1	PCT-US00-24802-3639	Sequence 3639, Ap
21	21	48.8	8 16	US-09-239-043-1134	Sequence 1134, Ap
22	21	48.8	8 16	US-09-239-043-1694	Sequence 1694, Ap
23	21	48.8	8 16	US-09-239-043-2441	Sequence 2441, Ap
24	21	48.8	8 16	US-09-239-043A-1134	Sequence 1134, Ap
25	21	48.8	8 16	US-09-239-043A-1694	Sequence 1694, Ap
26	21	48.8	8 16	US-09-239-043A-2441	Sequence 2441, Ap
27	21	48.8	8 17	US-09-350-401A-1169	Sequence 1169, Ap
28	21	48.8	8 17	US-09-350-401A-1661	Sequence 1661, Ap
29	21	48.8	8 17	US-09-350-401A-3341	Sequence 3341, Ap
30	21	48.8	8 17	US-09-350-401A-3639	Sequence 3639, Ap
31	21	48.8	8 26	US-10-030-014-472	Sequence 472, App
32	21	48.8	9 1	PCT-US00-24802-73	Sequence 73, Appl
33	21	48.8	9 1	PCT-US00-24802-222	Sequence 222, App
34	21	48.8	9 1	PCT-US00-24802-1420	Sequence 1420, Ap
35	21	48.8	9 1	PCT-US00-24802-1714	Sequence 1714, Ap
36	21	48.8	9 1	PCT-US00-24802-2146	Sequence 2146, Ap
37	21	48.8	9 1	PCT-US00-24802-2278	Sequence 2278, Ap
38	21	48.8	9 1	PCT-US00-24802-2964	Sequence 2964, Ap
39	21	48.8	9 1	PCT-US00-24802-3235	Sequence 3235, Ap
40	21	48.8	9 1	PCT-US00-24802-3505	Sequence 3505, Ap
41	21	48.8	9 1	PCT-US00-24802-3703	Sequence 3703, Ap
42	21	48.8	9 1	PCT-US00-28558-115	Sequence 115, App
43	21	48.8	9 1	PCT-US00-28558-121	Sequence 121, App
44	21	48.8	9 1	PCT-US00-28558-143	Sequence 143, App
45	21	48.8	9 1	PCT-US02-32727-30876	Sequence 30876, A

#### ALIGNMENTS

RESULT 1

US-09-807-512-24  
; Sequence 24, Application US/09807512  
; GENERAL INFORMATION:  
; APPLICANT: Schrier, Peter I.  
; APPLICANT: Aarnoudse, Corlien  
; APPLICANT: Heider, Karl-Heinz  
; APPLICANT: Klade, Christoph  
; TITLE OF INVENTION: Camel, An Alternative Translation Product of the Tumor  
; FILE REFERENCE: 0652.2200000  
; CURRENT APPLICATION NUMBER: US/09/807,512  
; CURRENT FILING DATE: 2001-04-16  
; PRIOR APPLICATION NUMBER: PCT/EP99/07832  
; PRIOR FILING DATE: 1999-10-15  
; PRIOR APPLICATION NUMBER: EP 98119583.7  
; PRIOR FILING DATE: 1998-10-16  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 24  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-807-512-24

Query Match 100.0%; Score 43; DB 23; Length 9;  
Best Local Similarity 100.0%; Pred. No. 5.3e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLMAQGAML 9  
Db 1 FLMAQGAML 9

RESULT 2  
US-09-641-377-1005  
; Sequence 1005, Application US/09641377  
; GENERAL INFORMATION:  
; APPLICANT: WIEMANN, STEFAN  
; APPLICANT: GASSENHUBER, JOHANN  
; APPLICANT: TAMPE, JENS  
; TITLE OF INVENTION: HUMAN DNA SEQUENCES  
; FILE REFERENCE: 087100/0106  
; CURRENT APPLICATION NUMBER: US/09/641,377  
; CURRENT FILING DATE: 2000-08-18  
; PRIOR APPLICATION NUMBER: 60/149,499  
; PRIOR FILING DATE: 1999-08-18  
; PRIOR APPLICATION NUMBER: 60/156,503  
; PRIOR FILING DATE: 1999-09-28  
; NUMBER OF SEQ ID NOS: 1793  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1005  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Consensus  
US-09-641-377-1005

Query Match 55.8%; Score 24; DB 20; Length 8;  
Best Local Similarity 100.0%; Pred. No. 5.3e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LMAQG 6  
Db 1 LMAQG 5

RESULT 3  
US-09-641-528A-3437  
; Sequence 3437, Application US/09641528A  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Grey, Howard  
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS  
; FILE REFERENCE: 2060.0100001  
; CURRENT APPLICATION NUMBER: US/09/641,528A  
; CURRENT FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: US 60/172,705  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 51504  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3437  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus  
US-09-641-528A-3437

Query Match 55.8%; Score 24; DB 20; Length 8;

Best Local Similarity 100.0%; Pred. No. 5.3e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 QGAML 9  
Db 2 QGAML 6

RESULT 4  
US-09-641-528A-18325  
; Sequence 18325, Application US/09641528A  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Grey, Howard  
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS  
; FILE REFERENCE: 2060.0100001  
; CURRENT APPLICATION NUMBER: US/09/641,528A  
; CURRENT FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: US 60/172,705  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 51504  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 18325  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus  
US-09-641-528A-18325

Query Match 55.8%; Score 24; DB 20; Length 8;  
Best Local Similarity 100.0%; Pred. No. 5.3e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 QGAML 9  
Db 3 QGAML 7

RESULT 5  
US-09-641-528A-25548  
; Sequence 25548, Application US/09641528A  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Grey, Howard  
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS  
; FILE REFERENCE: 2060.0100001  
; CURRENT APPLICATION NUMBER: US/09/641,528A  
; CURRENT FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: US 60/172,705  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 51504  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 25548  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus  
US-09-641-528A-25548

Query Match 55.8%; Score 24; DB 20; Length 8;  
Best Local Similarity 100.0%; Pred. No. 5.3e+06;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 QGAML 9  
|||||

Db 2 QGAML 6

RESULT 6  
US-09-641-528A-33046  
; Sequence 33046, Application US/09641528A  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Grey, Howard  
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS  
; FILE REFERENCE: 2060.0100001  
; CURRENT APPLICATION NUMBER: US/09/641,528A  
; CURRENT FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: US 60/172,705  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 51504  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 33046  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus  
US-09-641-528A-33046

Query Match 55.8%; Score 24; DB 20; Length 8;  
Best Local Similarity 100.0%; Pred. No. 5.3e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 QGAML 9  
|||||

Db 1 QGAML 5

RESULT 7  
US-09-641-528A-18326  
; Sequence 18326, Application US/09641528A  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Grey, Howard  
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS  
; FILE REFERENCE: 2060.0100001  
; CURRENT APPLICATION NUMBER: US/09/641,528A  
; CURRENT FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: US 60/172,705  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 51504  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 18326  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus  
US-09-641-528A-18326

Query Match 55.8%; Score 24; DB 20; Length 9;  
Best Local Similarity 100.0%; Pred. No. 5.3e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 QGAML 9  
|||||

Db 3 QGAML 7

RESULT 8  
US-09-641-528A-25549  
; Sequence 25549, Application US/09641528A  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Grey, Howard  
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS  
; FILE REFERENCE: 2060.0100001  
; CURRENT APPLICATION NUMBER: US/09/641,528A  
; CURRENT FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: US 60/172,705  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 51504  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 25549  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus  
US-09-641-528A-25549

Query Match 55.8%; Score 24; DB 20; Length 9;  
Best Local Similarity 100.0%; Pred. No. 5.3e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 QGAML 9  
|||||

Db 2 QGAML 6

RESULT 9  
US-09-641-528A-33047  
; Sequence 33047, Application US/09641528A  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Grey, Howard  
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS  
; FILE REFERENCE: 2060.0100001  
; CURRENT APPLICATION NUMBER: US/09/641,528A  
; CURRENT FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: US 60/172,705  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 51504  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 33047  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus  
US-09-641-528A-33047

Query Match 55.8%; Score 24; DB 20; Length 9;  
Best Local Similarity 100.0%; Pred. No. 5.3e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 QGAML 9  
| | | | |  
Db 1 QGAML 5

## RESULT 10

US-09-641-528A-40821  
; Sequence 40821, Application US/09641528A

; GENERAL INFORMATION:

; APPLICANT: Sette, Alessandro

; APPLICANT: Sidney, John

; APPLICANT: Southwood, Scott

; APPLICANT: Chesnut, Robert

; APPLICANT: Celis, Esteban

; APPLICANT: Grey, Howard

; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS

; FILE REFERENCE: 2060.0100001

; CURRENT APPLICATION NUMBER: US/09/641,528A

; CURRENT FILING DATE: 2000-08-15

; PRIOR FILING DATE: 1999-12-10

; NUMBER OF SEQ ID NOS: 51504

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 40821

; LENGTH: 9

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Peptide Derived from Human Papillomavirus

US-09-641-528A-40821

Query Match

Best Local Similarity 55.8%; Score 24; DB 20; Length 9;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 QGAML 9  
| | | | |  
Db 1 QGAML 5

## RESULT 11

US-09-932-165-266

; Sequence 266, Application US/09932165

; GENERAL INFORMATION:

; APPLICANT: RAITANO, ARTHUR

; APPLICANT: CHALLITA-EID, PIA M.

; APPLICANT: FARIS, MARY

; APPLICANT: SAFFRAN, DOUGLAS

; APPLICANT: AFAR, DANIEL

; APPLICANT: HUBERT, RENE

; APPLICANT: GE, WANGMAO

; APPLICANT: JAKOBOVITS, AYA

; TITLE OF INVENTION: NUCLEIC ACIDS AND CORRESPONDING PROTEINS ENTITLED

; TITLE OF INVENTION: 83PZH3 AND CATIFZELL USEFUL IN TREATMENT AND

; FILE REFERENCE: 51158-20014.00

; CURRENT APPLICATION NUMBER: US/09/932,165

; CURRENT FILING DATE: 2001-08-17

; PRIOR FILING DATE: 2000-08-17

; NUMBER OF SEQ ID NOS: 1508

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 266

; LENGTH: 9

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Peptide motif

US-09-932-165-266

Query Match

Best Local Similarity 55.8%; Score 24; DB 24; Length 9;

Best Local Similarity 83.3%; Pred. No. 5.3e+06;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LMAQGA 7  
| : | | | |  
Db 2 LVAQGA 7

## RESULT 12

US-10-182-252A-522

; Sequence 522, Application US/10182252A

; GENERAL INFORMATION:

; APPLICANT: FOMSGAARD, ANDERS

; APPLICANT: BRUNAK, SOREN

; APPLICANT: BUUS, SOREN

; APPLICANT: CORBET, SYLVIE

; APPLICANT: LAUEMOLLER, SANNE LISE

; APPLICANT: HANSEN, JAN

; TITLE OF INVENTION: HIV PEPTIDE AND NUCLEIC ACIDS ENCODING THEM FOR DIAGNOSIS AND

; FILE REFERENCE: 030307/0205

; CURRENT APPLICATION NUMBER: US/10/182,252A

; CURRENT FILING DATE: 2003-04-10

; PRIOR FILING DATE: 2001-01-29

; PRIOR APPLICATION NUMBER: EP 00610017.6

; PRIOR FILING DATE: 2000-01-28

; PRIOR APPLICATION NUMBER: US 60/179,333

; PRIOR FILING DATE: 2000-01-31

; NUMBER OF SEQ ID NOS: 1388

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 522

; LENGTH: 9

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: HIV peptide

US-10-182-252A-522

Query Match

Best Local Similarity 55.8%; Score 24; DB 27; Length 9;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FLMAQGA 9  
| | | | |  
Db 1 FLMIQGLL 9

## RESULT 13

US-10-031-874A-124

; Sequence 124, Application US/10031874A

; GENERAL INFORMATION:

; APPLICANT: TANHA, JAMSHID

; APPLICANT: DUBUC, GINETTE

; APPLICANT: NARANG, SARAN

; TITLE OF INVENTION: SINGLE-DOMAIN ANTIGEN-BINDING ANTIBODY FRAGMENTS

; TITLE OF INVENTION: DERIVED FROM LLAMA ANTIBODIES

; FILE REFERENCE: 11054-1

; CURRENT APPLICATION NUMBER: US/10/031,874A

; CURRENT FILING DATE: 2002-11-14

; PRIOR FILING DATE: 2000-05-26

; NUMBER OF SEQ ID NOS: 212

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 124

; LENGTH: 8

; TYPE: PRT

; ORGANISM: Lama glama

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Peptide motif

US-10-031-874A-124

Query Match

Best Local Similarity 53.5%; Score 23; DB 26; Length 8;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 MAQGAM 8  
:::||||  
Db 1 LSQGAM 6

RESULT 14  
US-10-031-874A-125  
; Sequence 125, Application US/10031874A  
; GENERAL INFORMATION:  
; APPLICANT: TANHA, JAMSHID  
; APPLICANT: DUBUC, GINETTE  
; APPLICANT: NARANG, SARAN  
; TITLE OF INVENTION: SINGLE-DOMAIN ANTIGEN-BINDING ANTIBODY FRAGMENTS  
; FILE REFERENCE: 11054-1  
; CURRENT APPLICATION NUMBER: US/10/031,874A  
; CURRENT FILING DATE: 2002-11-14  
; PRIOR APPLICATION NUMBER: 60/207,234  
; PRIOR FILING DATE: 2000-05-26  
; NUMBER OF SEQ ID NOS: 212  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 125  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Lama glama  
US-10-031-874A-125

Query Match 53.5%; Score 23; DB 26; Length 8;  
Best Local Similarity 66.7%; Pred. No. 5.3e+06;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 MAQGAM 8  
:::||||  
Db 1 LSQGAM 6

RESULT 15  
US-10-031-874A-126  
; Sequence 126, Application US/10031874A  
; GENERAL INFORMATION:  
; APPLICANT: TANHA, JAMSHID  
; APPLICANT: DUBUC, GINETTE  
; APPLICANT: NARANG, SARAN  
; TITLE OF INVENTION: SINGLE-DOMAIN ANTIGEN-BINDING ANTIBODY FRAGMENTS  
; FILE REFERENCE: 11054-1  
; CURRENT APPLICATION NUMBER: US/10/031,874A  
; CURRENT FILING DATE: 2002-11-14  
; PRIOR APPLICATION NUMBER: 60/207,234  
; PRIOR FILING DATE: 2000-05-26  
; NUMBER OF SEQ ID NOS: 212  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 126  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Lama glama  
US-10-031-874A-126

Query Match 53.5%; Score 23; DB 26; Length 8;  
Best Local Similarity 66.7%; Pred. No. 5.3e+06;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 MAQGAM 8  
:::||||  
Db 1 LSQGAM 6

Search completed: October 7, 2003, 15:00:03  
Job time : 278.5 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:48:22 ; Search time 14.5 seconds  
(without alignments)  
18.078 Million cell updates/sec

Title: US-09-807-512-24  
Perfect score: 43  
Sequence: 1 FLMAQGAWL 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 136899 seqs, 29126274 residues

Total number of hits satisfying chosen parameters: 17417

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*  
1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US05\_NEW\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	21	48.8	8	5 US-09-239-043D-1134	Sequence 1134, Ap
2	21	48.8	8	5 US-09-239-043D-1694	Sequence 1694, Ap
3	21	48.8	8	5 US-09-239-043D-2441	Sequence 2441, Ap
4	21	48.8	8	6 US-10-363-990-1169	Sequence 1169, Ap
5	21	48.8	8	6 US-10-363-990-1661	Sequence 1661, Ap
6	21	48.8	8	6 US-10-363-990-3341	Sequence 3341, Ap
7	21	48.8	8	6 US-10-363-990-3639	Sequence 3639, Ap
8	21	48.8	9	5 US-09-239-043D-73	Sequence 73, Appl
9	21	48.8	9	5 US-09-239-043D-214	Sequence 214, Ap
10	21	48.8	9	5 US-09-239-043D-1401	Sequence 1401, Ap
11	21	48.8	9	5 US-09-239-043D-1869	Sequence 1869, Ap
12	21	48.8	9	5 US-09-239-043D-2270	Sequence 2270, Ap
13	21	48.8	9	6 US-10-363-990-73	Sequence 73, Appl
14	21	48.8	9	6 US-10-363-990-222	Sequence 222, Ap
15	21	48.8	9	6 US-10-363-990-1420	Sequence 1420, Ap
16	21	48.8	9	6 US-10-363-990-1714	Sequence 1714, Ap
17	21	48.8	9	6 US-10-363-990-2146	Sequence 2146, Ap
18	21	48.8	9	6 US-10-363-990-2278	Sequence 2278, Ap
19	21	48.8	9	6 US-10-363-990-2964	Sequence 2964, Ap
20	21	48.8	9	6 US-10-363-990-3235	Sequence 3235, Ap
21	21	48.8	9	6 US-10-363-990-3505	Sequence 3505, Ap
22	21	48.8	9	6 US-10-363-990-3703	Sequence 3703, Ap
23	20	46.5	9	1 PCT-US03-11231-263	Sequence 263, Appl
24	20	46.5	9	7 US-60-490-788-26	Sequence 26, Appl
25	19	44.2	7	4 US-08-823-980F-4	Sequence 4, Appl
26	19	44.2	9	6 US-10-647-005-15	Sequence 15, Appl

Sequence 8, Appli  
Sequence 113, App  
Sequence 125, App  
Sequence 227, App  
Sequence 240, App  
Sequence 430, App  
Sequence 440, App  
Sequence 510, App  
Sequence 549, App  
Sequence 549, App  
Sequence 644, App  
Sequence 244, App  
Sequence 216, App  
Sequence 716, App  
Sequence 2527, Ap  
Sequence 224, App  
Sequence 751, App  
Sequence 1715, Ap  
Sequence 3507, Ap  
Sequence 3705, Ap

ALIGNMENTS

RESULT 1  
US-09-239-043D-1134  
; Sequence 1134, Application US/09239043D  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Vitello, Maria A.  
; APPLICANT: Livingston, Brian D.  
; APPLICANT: Celis, Esteban  
; APPLICANT: Kubo, Ralph T.  
; APPLICANT: Grey, Howard M.  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Epimmune Inc.  
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Hepatitis B Virus  
; TITLE OF INVENTION: Using Peptide and Nucleic Acid Compositions  
; FILE REFERENCE: 2060.0060007  
; CURRENT APPLICATION NUMBER: US/09/239,043D  
; CURRENT FILING DATE: 1999-01-27  
; PRIOR APPLICATION NUMBER: US 09/189,702  
; PRIOR FILING DATE: 1998-11-10  
; PRIOR APPLICATION NUMBER: US 08/978,291  
; PRIOR FILING DATE: 1997-11-25  
; PRIOR APPLICATION NUMBER: US 08/820,360  
; PRIOR FILING DATE: 1997-03-12  
; PRIOR APPLICATION NUMBER: US 60/013,363  
; PRIOR FILING DATE: 1996-03-13  
; PRIOR APPLICATION NUMBER: US 08/461,603  
; PRIOR FILING DATE: 1995-06-05  
; PRIOR APPLICATION NUMBER: US 08/347,610  
; PRIOR FILING DATE: 1994-12-01  
; PRIOR APPLICATION NUMBER: US 08/344,824  
; PRIOR FILING DATE: 1994-11-23  
; PRIOR APPLICATION NUMBER: US 08/278,634  
; PRIOR FILING DATE: 1994-07-21  
; PRIOR APPLICATION NUMBER: US 08/205,713  
; PRIOR FILING DATE: 1994-03-04  
; PRIOR APPLICATION NUMBER: US 08/197,484  
; PRIOR FILING DATE: 1994-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 2579  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1134  
; LENGTH: 8  
; TYPE: PPT  
; ORGANISM: Orthohepadnaviridae hepatitis B virus  
US-09-239-043D-1134

Query Match 48.8%; Score 21; DB 5; Length 8;

```
Best Local Similarity 80.0%; Pred. No. 1.2e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAQ 5
Db 3 FLIAQ 7

RESULT 2
US-09-239-043D-1694
; Sequence 1694, Application US/09239043D
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Vitiello, Maria A.
; APPLICANT: Livingston, Brian D.
; APPLICANT: Celis, Esteban
; APPLICANT: Kubo, Ralph T.
; APPLICANT: Grey, Howard M.
; APPLICANT: Chesnut, Robert
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Hepatitis B Virus
; FILE REFERENCE: 2060.0060007
; CURRENT APPLICATION NUMBER: US/09/239,043D
; CURRENT FILING DATE: 1999-01-27
; PRIOR APPLICATION NUMBER: US 09/189,702
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: US 08/978,291
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: US 08/820,360
; PRIOR FILING DATE: 1997-03-12
; PRIOR APPLICATION NUMBER: US 60/013,363
; PRIOR FILING DATE: 1996-03-13
; PRIOR APPLICATION NUMBER: US 08/461,603
; PRIOR FILING DATE: 1995-06-05
; PRIOR APPLICATION NUMBER: US 08/347,610
; PRIOR FILING DATE: 1994-12-01
; PRIOR APPLICATION NUMBER: US 08/344,824
; PRIOR FILING DATE: 1994-11-23
; PRIOR APPLICATION NUMBER: US 08/278,634
; PRIOR FILING DATE: 1994-07-21
; PRIOR APPLICATION NUMBER: US 08/205,713
; PRIOR FILING DATE: 1994-03-04
; PRIOR APPLICATION NUMBER: US 08/197,484
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2579
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1694
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Orthohepadnaviridae hepatitis B virus
; OTHER INFORMATION: HBV analog peptide
US-09-239-043D-1694

Query Match 48.8%; Score 21; DB 5; Length 8;
Best Local Similarity 80.0%; Pred. No. 1.2e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAQ 5
Db 3 FLIAQ 7

RESULT 3
US-09-239-043D-2441
; Sequence 2441, Application US/09239043D
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Vitiello, Maria A.
; APPLICANT: Chesnut, Robert
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Hepatitis B
; FILE REFERENCE: 2060.0060009
; CURRENT APPLICATION NUMBER: US/10/363,990
; CURRENT FILING DATE: 2003-03-10
; PRIOR APPLICATION NUMBER: WO PCT/US00/24802
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 3876
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
APPLICANT: Livingston, Brian D.
APPLICANT: Celis, Esteban
APPLICANT: Kubo, Ralph T.
APPLICANT: Grey, Howard M.
APPLICANT: Chesnut, Robert
APPLICANT: Epimmune Inc.
TITLE OF INVENTION: Inducing Cellular Immune Responses to Hepatitis B Virus
FILE REFERENCE: 2060.0060007
CURRENT APPLICATION NUMBER: US/09/239,043D
CURRENT FILING DATE: 1999-01-27
PRIOR APPLICATION NUMBER: US 09/189,702
PRIOR FILING DATE: 1998-11-10
PRIOR APPLICATION NUMBER: US 08/978,291
PRIOR FILING DATE: 1997-11-25
PRIOR APPLICATION NUMBER: US 08/820,360
PRIOR FILING DATE: 1997-03-12
PRIOR APPLICATION NUMBER: US 60/013,363
PRIOR FILING DATE: 1996-03-13
PRIOR APPLICATION NUMBER: US 08/461,603
PRIOR FILING DATE: 1995-06-05
PRIOR APPLICATION NUMBER: US 08/347,610
PRIOR FILING DATE: 1994-12-01
PRIOR APPLICATION NUMBER: US 08/344,824
PRIOR FILING DATE: 1994-11-23
PRIOR APPLICATION NUMBER: US 08/278,634
PRIOR FILING DATE: 1994-07-21
PRIOR APPLICATION NUMBER: US 08/205,713
PRIOR FILING DATE: 1994-03-04
PRIOR APPLICATION NUMBER: US 08/197,484
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 2579
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2441
LENGTH: 8
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: HBV analog peptide
US-09-239-043D-2441

Query Match 48.8%; Score 21; DB 5; Length 8;
Best Local Similarity 80.0%; Pred. No. 1.2e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAQ 5
Db 3 FLIAQ 7

RESULT 4
US-10-363-990-1169
; Sequence 1169, Application US/10363990
; GENERAL INFORMATION:
; APPLICANT: Alessandro Sette
; APPLICANT: John Sidney
; APPLICANT: Scott Southwood
; APPLICANT: Maria A. Vitiello
; APPLICANT: Brian D. Livingston
; APPLICANT: Esteban Celis
; APPLICANT: Ralph T. Kubo
; APPLICANT: Howard M. Grey
; APPLICANT: Robert Chesnut
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Hepatitis B
; FILE REFERENCE: 2060.0060009
; CURRENT APPLICATION NUMBER: US/10/363,990
; CURRENT FILING DATE: 2003-03-10
; PRIOR APPLICATION NUMBER: WO PCT/US00/24802
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 3876
; SOFTWARE: FastSeq for Windows Version 4.0
```

; SEQ ID NO 1169  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Artificially Synthesized Peptide  
US-10-363-990-1169

Query Match 48.8%; Score 21; DB 6; Length 8;  
Best Local Similarity 80.0%; Pred. No. 1.2e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAQ 5  
||:|  
Db 3 FLAQ 7

RESULT 5  
US-10-363-990-1661  
; Sequence 1661, Application US/10363990  
; GENERAL INFORMATION:  
; APPLICANT: Alessandro Sette  
; APPLICANT: John Sidney  
; APPLICANT: Scott Southwood  
; APPLICANT: Maria A. Vitello  
; APPLICANT: Brian D. Livingston  
; APPLICANT: Esteban Celis  
; APPLICANT: Ralph T. Kubo  
; APPLICANT: Howard M. Grey  
; APPLICANT: Robert Chesnut  
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Hepatitis B  
; FILE REFERENCE: 2060.0060009  
; CURRENT APPLICATION NUMBER: US/10/363,990  
; CURRENT FILING DATE: 2003-03-10  
; PRIOR APPLICATION NUMBER: WO PCT/US00/24802  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 3876  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1661  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Artificially Synthesized Peptide  
US-10-363-990-1661

Query Match 48.8%; Score 21; DB 6; Length 8;  
Best Local Similarity 80.0%; Pred. No. 1.2e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAQ 5  
||:|  
Db 3 FLAQ 7

RESULT 6  
US-10-363-990-3341  
; Sequence 3341, Application US/10363990  
; GENERAL INFORMATION:  
; APPLICANT: Alessandro Sette  
; APPLICANT: John Sidney  
; APPLICANT: Scott Southwood  
; APPLICANT: Maria A. Vitello  
; APPLICANT: Brian D. Livingston  
; APPLICANT: Esteban Celis  
; APPLICANT: Ralph T. Kubo  
; APPLICANT: Howard M. Grey  
; APPLICANT: Robert Chesnut  
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Hepatitis B  
; FILE REFERENCE: 2060.0060009  
; CURRENT APPLICATION NUMBER: US/10/363,990

; CURRENT FILING DATE: 2003-03-10  
; PRIOR APPLICATION NUMBER: WO PCT/US00/24802  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 3876  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3341  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Artificially Synthesized Peptide  
US-10-363-990-3341

Query Match 48.8%; Score 21; DB 6; Length 8;  
Best Local Similarity 80.0%; Pred. No. 1.2e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAQ 5  
||:|  
Db 3 FLAQ 7

RESULT 7  
US-10-363-990-3639  
; Sequence 3639, Application US/10363990  
; GENERAL INFORMATION:  
; APPLICANT: Alessandro Sette  
; APPLICANT: John Sidney  
; APPLICANT: Scott Southwood  
; APPLICANT: Maria A. Vitello  
; APPLICANT: Brian D. Livingston  
; APPLICANT: Esteban Celis  
; APPLICANT: Ralph T. Kubo  
; APPLICANT: Howard M. Grey  
; APPLICANT: Robert Chesnut  
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Hepatitis B  
; FILE REFERENCE: 2060.0060009  
; CURRENT APPLICATION NUMBER: US/10/363,990  
; CURRENT FILING DATE: 2003-03-10  
; PRIOR APPLICATION NUMBER: WO PCT/US00/24802  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 3876  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3639  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Artificial Peptide  
US-10-363-990-3639

Query Match 48.8%; Score 21; DB 6; Length 8;  
Best Local Similarity 80.0%; Pred. No. 1.2e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAQ 5  
||:|  
Db 3 FLAQ 7

RESULT 8  
US-09-239-043D-73  
; Sequence 73, Application US/09239043D  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Vitello, Maria A.  
; APPLICANT: Livingston, Brian D.  
; APPLICANT: Celis, Esteban  
; APPLICANT: Kubo, Ralph T.  
; APPLICANT: Grey, Howard M.

```
; APPLICANT: Chesnut, Robert
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Hepatitis B Virus
; FILE REFERENCE: 2060.0060007
; CURRENT APPLICATION NUMBER: US/09/239,043D
; PRIOR FILING DATE: 1999-01-27
; PRIOR APPLICATION NUMBER: US 09/189,702
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: US 08/978,291
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: US 08/820,360
; PRIOR FILING DATE: 1997-03-12
; PRIOR APPLICATION NUMBER: US 60/013,363
; PRIOR FILING DATE: 1996-03-13
; PRIOR APPLICATION NUMBER: US 08/461,603
; PRIOR FILING DATE: 1995-06-05
; PRIOR APPLICATION NUMBER: US 08/347,610
; PRIOR FILING DATE: 1994-12-01
; PRIOR APPLICATION NUMBER: US 08/344,824
; PRIOR FILING DATE: 1994-11-23
; PRIOR APPLICATION NUMBER: US 08/278,634
; PRIOR FILING DATE: 1994-07-21
; PRIOR APPLICATION NUMBER: US 08/205,713
; PRIOR FILING DATE: 1994-03-04
; PRIOR APPLICATION NUMBER: US 08/197,484
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2579
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Orthohepadnaviridae hepatitis B virus
US-09-239-043D-73

Query Match      48.8%; Score 21; DB 5; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.2e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAQ 5
Db 4 FLLAQ 8

RESULT 9
US-09-239-043D-214
; Sequence 214, Application US/09239043D
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Vitiello, Maria A.
; APPLICANT: Livingston, Brian D.
; APPLICANT: Celis, Esteban
; APPLICANT: Kubo, Ralph T.
; APPLICANT: Grey, Howard M.
; APPLICANT: Chesnut, Robert
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Hepatitis B Virus
; FILE REFERENCE: 2060.0060007
; CURRENT APPLICATION NUMBER: US/09/239,043D
; CURRENT FILING DATE: 1999-01-27
; PRIOR APPLICATION NUMBER: US 09/189,702
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: US 08/978,291
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: US 08/820,360
; PRIOR FILING DATE: 1997-03-12
; PRIOR APPLICATION NUMBER: US 60/013,363
; PRIOR FILING DATE: 1996-03-13
; PRIOR APPLICATION NUMBER: US 08/461,603
; PRIOR FILING DATE: 1995-06-05
; PRIOR APPLICATION NUMBER: US 08/347,610
; PRIOR FILING DATE: 1994-12-01
; PRIOR APPLICATION NUMBER: US 08/344,824
; PRIOR FILING DATE: 1994-11-23
; PRIOR APPLICATION NUMBER: US 08/278,634
; PRIOR FILING DATE: 1994-07-21
; PRIOR APPLICATION NUMBER: US 08/205,713
; PRIOR FILING DATE: 1994-03-12
; PRIOR APPLICATION NUMBER: US 60/013,363
; PRIOR FILING DATE: 1996-03-13
; PRIOR APPLICATION NUMBER: US 08/461,603
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2579
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Orthohepadnaviridae hepatitis B virus
US-09-239-043D-73

Query Match      48.8%; Score 21; DB 5; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.2e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAQ 5
Db 4 FLLAQ 8

RESULT 9
US-09-239-043D-214
; Sequence 214, Application US/09239043D
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Vitiello, Maria A.
; APPLICANT: Livingston, Brian D.
; APPLICANT: Celis, Esteban
; APPLICANT: Kubo, Ralph T.
; APPLICANT: Grey, Howard M.
; APPLICANT: Chesnut, Robert
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Hepatitis B Virus
; FILE REFERENCE: 2060.0060007
; CURRENT APPLICATION NUMBER: US/09/239,043D
; CURRENT FILING DATE: 1999-01-27
; PRIOR APPLICATION NUMBER: US 09/189,702
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: US 08/978,291
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: US 08/820,360
; PRIOR FILING DATE: 1997-03-12
; PRIOR APPLICATION NUMBER: US 60/013,363
; PRIOR FILING DATE: 1996-03-13
; PRIOR APPLICATION NUMBER: US 08/461,603
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2579
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Orthohepadnaviridae hepatitis B virus
US-09-239-043D-73
```

```
; PRIOR FILING DATE: 1995-06-05
; PRIOR APPLICATION NUMBER: US 08/347,610
; PRIOR FILING DATE: 1994-12-01
; PRIOR APPLICATION NUMBER: US 08/344,824
; PRIOR FILING DATE: 1994-11-23
; PRIOR APPLICATION NUMBER: US 08/278,634
; PRIOR FILING DATE: 1994-07-21
; PRIOR APPLICATION NUMBER: US 08/205,713
; PRIOR FILING DATE: 1994-03-04
; PRIOR APPLICATION NUMBER: US 08/197,484
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2579
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 214
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Orthohepadnaviridae hepatitis B virus
US-09-239-043D-214

Query Match      48.8%; Score 21; DB 5; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.2e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAQ 5
Db 1 FLLAQ 5

RESULT 10
US-09-239-043D-1401
; Sequence 1401, Application US/09239043D
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Vitiello, Maria A.
; APPLICANT: Livingston, Brian D.
; APPLICANT: Celis, Esteban
; APPLICANT: Kubo, Ralph T.
; APPLICANT: Grey, Howard M.
; APPLICANT: Chesnut, Robert
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Hepatitis B Virus
; FILE REFERENCE: 2060.0060007
; CURRENT APPLICATION NUMBER: US/09/239,043D
; CURRENT FILING DATE: 1999-01-27
; PRIOR APPLICATION NUMBER: US 09/189,702
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: US 08/978,291
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: US 08/820,360
; PRIOR FILING DATE: 1997-03-12
; PRIOR APPLICATION NUMBER: US 60/013,363
; PRIOR FILING DATE: 1996-03-13
; PRIOR APPLICATION NUMBER: US 08/461,603
; PRIOR FILING DATE: 1995-06-05
; PRIOR APPLICATION NUMBER: US 08/347,610
; PRIOR FILING DATE: 1994-12-01
; PRIOR APPLICATION NUMBER: US 08/344,824
; PRIOR FILING DATE: 1994-11-23
; PRIOR APPLICATION NUMBER: US 08/278,634
; PRIOR FILING DATE: 1994-07-21
; PRIOR APPLICATION NUMBER: US 08/205,713
; PRIOR FILING DATE: 1994-03-04
; PRIOR APPLICATION NUMBER: US 08/197,484
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2579
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1401
; LENGTH: 9
```

```
; TYPE: PRT
; ORGANISM: Orthohepadnaviridae hepatitis B virus
US-09-239-043D-1401

Query Match          48.8%; Score 21; DB 5; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.2e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAQ 5
   ||||
Db 4 FLAQ 8

RESULT 11
US-09-239-043D-1869
; Sequence 1869, Application US/09239043D
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Vitiello, Maria A.
; APPLICANT: Livingston, Brian D.
; APPLICANT: Celis, Esteban
; APPLICANT: Kubo, Ralph T.
; APPLICANT: Grey, Howard M.
; APPLICANT: Chesnut, Robert
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Hepatitis B Virus
; FILE REFERENCE: 2060.0060007
; CURRENT APPLICATION NUMBER: US/09/239,043D
; CURRENT FILING DATE: 1999-01-27
; PRIOR APPLICATION NUMBER: US 09/189,702
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: US 08/978,291
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: US 08/820,360
; PRIOR FILING DATE: 1997-03-12
; PRIOR APPLICATION NUMBER: US 60/013,363
; PRIOR FILING DATE: 1996-03-13
; PRIOR APPLICATION NUMBER: US 08/461,603
; PRIOR FILING DATE: 1995-06-05
; PRIOR APPLICATION NUMBER: US 08/347,610
; PRIOR FILING DATE: 1994-12-01
; PRIOR APPLICATION NUMBER: US 08/344,824
; PRIOR FILING DATE: 1994-11-23
; PRIOR APPLICATION NUMBER: US 08/278,634
; PRIOR FILING DATE: 1994-07-21
; PRIOR APPLICATION NUMBER: US 08/205,713
; PRIOR FILING DATE: 1994-03-04
; PRIOR APPLICATION NUMBER: US 08/197,484
; PRIOR FILING DATE: 1994-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2579
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1869
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Orthohepadnaviridae hepatitis B virus
US-09-239-043D-1869

Query Match          48.8%; Score 21; DB 5; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.2e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAQ 5
   ||||
Db 4 FLAQ 8

RESULT 12
US-09-239-043D-2270
; Sequence 2270, Application US/09239043D
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Vitiello, Maria A.
; APPLICANT: Livingston, Brian D.
; APPLICANT: Celis, Esteban
; APPLICANT: Kubo, Ralph T.
; APPLICANT: Grey, Howard M.
; APPLICANT: Chesnut, Robert
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Hepatitis B Virus
; FILE REFERENCE: 2060.0060007
; CURRENT APPLICATION NUMBER: US/09/239,043D
; CURRENT FILING DATE: 1999-01-27
; PRIOR APPLICATION NUMBER: US 09/189,702
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: US 08/978,291
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: US 08/820,360
; PRIOR FILING DATE: 1997-03-12
; PRIOR APPLICATION NUMBER: US 60/013,363
; PRIOR FILING DATE: 1996-03-13
; PRIOR APPLICATION NUMBER: US 08/461,603
; PRIOR FILING DATE: 1995-06-05
; PRIOR APPLICATION NUMBER: US 08/347,610
; PRIOR FILING DATE: 1994-12-01
; PRIOR APPLICATION NUMBER: US 08/344,824
; PRIOR FILING DATE: 1994-11-23
; PRIOR APPLICATION NUMBER: US 08/278,634
; PRIOR FILING DATE: 1994-07-21
; PRIOR APPLICATION NUMBER: US 08/205,713
; PRIOR FILING DATE: 1994-03-04
; PRIOR APPLICATION NUMBER: US 08/197,484
; PRIOR FILING DATE: 1994-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2579
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1869
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Orthohepadnaviridae hepatitis B virus
US-09-239-043D-1869

Query Match          48.8%; Score 21; DB 5; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.2e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMAQ 5
   ||||
Db 4 FLAQ 8

RESULT 13
US-10-363-990-73
; Sequence 73, Application US/10363990
; GENERAL INFORMATION:
; APPLICANT: Alessandro Sette
; APPLICANT: John Sidney
; APPLICANT: Scott Southwood
; APPLICANT: Maria A. Vitiello
; APPLICANT: Brian D. Livingston
; APPLICANT: Esteban Celis
; APPLICANT: Ralph T. Kubo
; APPLICANT: Howard M. Grey
; APPLICANT: Robert Chesnut
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Hepatitis B
; FILE REFERENCE: 2060.0060009
; CURRENT APPLICATION NUMBER: US/10/363,990
```

```

; CURRENT FILING DATE: 2003-03-10
; PRIOR APPLICATION NUMBER: WO PCT/US00/24802
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 3876
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificially Synthesized Peptide
US-10-363-990-73

```

```

Query Match      48.8%; Score 21; DB 6; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.2e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 FLMAQ 5
        ||:|
Db      4 FLMAQ 8

```

```

RESULT 14
US-10-363-990-222
; Sequence 222, Application US/10363990
; GENERAL INFORMATION:
; APPLICANT: Alessandro Sette
; APPLICANT: John Sidney
; APPLICANT: Scott Southwood
; APPLICANT: Maria A. Vitiello
; APPLICANT: Brian D. Livingston
; APPLICANT: Esteban Celis
; APPLICANT: Ralph T. Kubo
; APPLICANT: Howard M. Grey
; APPLICANT: Robert Chesnut
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Hepatitis B
; FILE REFERENCE: 2060.0060009
; CURRENT APPLICATION NUMBER: US/10/363,990
; CURRENT FILING DATE: 2003-03-10
; PRIOR APPLICATION NUMBER: WO PCT/US00/24802
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 3876
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 222
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificially Synthesized Peptide
US-10-363-990-222

```

```

Query Match      48.8%; Score 21; DB 6; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.2e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 FLMAQ 5
        ||:|
Db      1 FLMAQ 5

```

```

RESULT 15
US-10-363-990-1420
; Sequence 1420, Application US/10363990
; GENERAL INFORMATION:
; APPLICANT: Alessandro Sette
; APPLICANT: John Sidney
; APPLICANT: Scott Southwood
; APPLICANT: Maria A. Vitiello
; APPLICANT: Brian D. Livingston
; APPLICANT: Esteban Celis
; APPLICANT: Ralph T. Kubo
; APPLICANT: Howard M. Grey

```

```

; APPLICANT: Robert Chesnut
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Hepatitis B
; FILE REFERENCE: 2060.0060009
; CURRENT APPLICATION NUMBER: US/10/363,990
; CURRENT FILING DATE: 2003-03-10
; PRIOR APPLICATION NUMBER: WO PCT/US00/24802
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 3876
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1420
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificially Synthesized Peptide
US-10-363-990-1420

```

```

Query Match      48.8%; Score 21; DB 6; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.2e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 FLMAQ 5
        ||:|
Db      4 FLMAQ 8

```

```

Search completed: October 7, 2003, 15:00:38
Job time : 15 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:16:45 ; Search time 57.5 seconds  
(without alignments)  
27.605 Million cell updates/sec

Title: US-09-807-512-25

Perfect score: 45

Sequence: 1 AMLAQERRV 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 251420

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	100.0	10	21	AA1980
2	41	91.1	9	21	AA1981
3	36	80.0	10	22	AA1982
4	36	80.0	10	22	AA1983
5	36	80.0	10	23	AA1984
6	32	71.1	9	20	AA1985
7	26	57.8	10	23	AA1986
8	25	55.6	9	21	AA1987
9	25	55.6	9	23	AA1988

10	24	53.3	9	24	ABP75370	Chlamydia trachoma
11	23	51.1	9	22	AA62638	Hsp47 HLA-A2 homol
12	23	51.1	10	22	AA697495	Human Complementar
13	22	48.9	7	21	AA607254	Immunomodulatory p
14	22	48.9	7	23	AAE20692	Immunomodulatory p
15	22	48.9	8	14	AA36590	Immunodepressed st
16	22	48.9	8	21	AA607255	Immunomodulatory p
17	22	48.9	8	23	AAE20693	Immunomodulatory p
18	22	48.9	9	21	AA609857	Immunomodulatory p
19	22	48.9	9	22	AA69558	Human Repro-EN-1.0
20	22	48.9	9	23	AAE20694	Immunomodulatory p
21	22	48.9	10	22	AA695984	Human Complementar
22	22	48.9	10	23	ABG79072	Human CAMEL class
23	22	48.9	10	23	AAU93544	Granulocyte-colony
24	22	48.9	10	24	ABR18923	Human cancer-relat
25	22	48.9	10	24	ABR19080	Human cancer-relat
26	21	46.7	6	20	AA23550	T-cell receptor V
27	21	46.7	7	22	AA80661	Human glandular ka
28	21	46.7	7	23	AAU85848	Prostate-specific
29	21	46.7	8	22	AA80698	Human glandular ka
30	21	46.7	9	20	AA33516	Human p75NTR depen
31	21	46.7	9	20	AA45834	Immunogenic peptid
32	21	46.7	9	20	AAW97915	Human synaptosomal
33	21	46.7	9	21	AA33294	HLA A1 binding TAD
34	21	46.7	9	21	AA33211	HLA A24 binding TA
35	21	46.7	9	21	AA33262	HLA B4403 binding
36	21	46.7	9	21	AA81906	Hsp47-related immu
37	21	46.7	9	21	AA76653	SCP-1 HLA binding
38	21	46.7	9	22	AAU68976	Human TADG-12 immu
39	21	46.7	9	22	AAU68993	Human TADG-12 immu
40	21	46.7	9	22	AAU69044	Human TADG-12 immu
41	21	46.7	9	22	AA898549	Human TADG-15 pept
42	21	46.7	10	20	AAW97925	Human synaptosomal
43	21	46.7	10	20	AAW99192	G protein alpha su
44	21	46.7	10	21	AA76663	SCP-1 HLA binding
45	21	46.7	10	22	AA83829	Arabidopsis thalia

#### ALIGNMENTS

##### RESULT 1

AA198058  
ID AA198058 standard; peptide; 10 AA.

XX  
AC AA198058;

XX  
XX 31-JUL-2000 (first entry)

XX  
DE CAMEL16 immunogenic peptide of human CAMEL protein.

XX  
DE CAMEL: CTL-recognised Antigen on MELANOMA: cytotoxic T lymphocyte; CTL;  
KW tumour-associated antigen; LAGE-1; NY-ESO-1; anticancer; melanoma;  
KW human; cancer; immunotherapy; immunogenic peptide; immune response.

XX  
OS Homo sapiens.

XX  
PN WO200023584-A1.

XX  
PD 27-APR-2000.

XX  
PF 15-OCT-1999; 99WO-EP07832.

XX  
PR 16-OCT-1998; 98EP-0119583.

XX  
PA (BOEH ) BOEHRINGER INGELHEIM INT GMBH.

XX  
PA (UYHO-) UNIV HOSPITAL LEIDEN.

XX  
PI Schrier PI, Aarnoudse CA, Heider K, Klade C;

XX  
XX WPI; 2000-339685/29.

XX  
PT Tumor-associated antigen useful for cancer immunotherapy is encoded by

PT the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA -

PS Claim 7; Page 34; 73pp; English.

XX The present sequence is an immunogenic peptide CAMEL 16, of the human

CC tumor-associated antigen CAMEL (Cytotoxic T lymphocytes (CTL)-recognised

CC Antigen on Melanoma). This peptide has the potential to bind to HLA-A2

CC and corresponds to residues 16-25 of the CAMEL protein. The CAMEL protein

CC is encoded by the LAGE-1 gene, a tumour-specific antigen. It is different

CC from the LAGE-1 protein, since it is translated from a different open

CC reading frame (ORF-1). It shows strong homology with NY-ESO-1, a melanoma

CC specific tumour antigen. The tumour-associated antigen displayed on

CC melanoma cells is recognised by cytotoxic T lymphocytes. CAMEL is

CC expressed in tumour cell lines, tumour tissues (e.g. breast and lung) and

CC in restricted number of healthy tissues. This sequence has anticancer

CC activity. CAMEL tumour antigen and immunogenic peptides derived from it

CC are useful for cancer immunotherapy. They have the potential to induce an

CC immune response, by eliciting a CTL response. The DNA molecule is used to

CC construct recombinant or fusion proteins.

XX

SO Sequence 10 AA;

Query Match 100.0%; Score 45; DB 21; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.025;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AMLAAQERRV 10

Db 1 AMLAAQERRV 10

|||||

RESULT 2

AAAY70859

ID AAY70859 standard; peptide; 9 AA.

XX

AC AAY70859;

XX

DT 31-JUL-2000 (first entry)

DE

DE CAMEL17 immunogenic peptide of human CAMEL protein.

XX

XX CAMEL; CTL-recognised Antigen on Melanoma; cytotoxic T lymphocyte; CTL;

KW tumour-associated antigen; LAGE-1; NY-ESO-1; anticancer; melanoma;

KW human; cancer; immunotherapy; immunogenic peptide; immune response.

XX

OS Homo sapiens.

XX

PN WO200023584-A1.

XX

PD 27-APR-2000.

XX

PF 15-OCT-1999; 99WO-EP07832.

XX

PR 16-OCT-1998; 98EP-0119583.

XX

PA (BOEH ) BOEHRINGER INGELHEIM INT GMBH.

PA (UYHO-) UNIV HOSPITAL LEIDEN.

XX

PI Schrier PI, Aarnoudse CA, Heider K, Klade C;

XX

XX WPI; 2000-339685/29.

XX

XX Tumor-associated antigen useful for cancer immunotherapy is encoded by

PT the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA -

XX

PS Claim 8; Page 34; 73pp; English.

XX

XX The present sequence is an immunogenic peptide CAMEL 17, of the human

CC tumor-associated antigen CAMEL (Cytotoxic T lymphocytes (CTL)-recognised

CC Antigen on Melanoma). This peptide has the potential to bind to HLA-A2

CC and corresponds to residues 17-25 of the CAMEL protein. The CAMEL protein

CC is encoded by the LAGE-1 gene, a tumour-specific antigen. It is different

CC from the LAGE-1 protein, since it is translated from a different open

CC reading frame (ORF-1). It shows strong homology with NY-ESO-1, a melanoma

CC specific tumour antigen. The tumour-associated antigen displayed on

CC melanoma cells is recognised by cytotoxic T lymphocytes. CAMEL is

CC - expressed in tumour cell lines, tumour tissues (e.g. breast and lung) and

CC in restricted number of healthy tissues. This sequence has anticancer

CC activity. CAMEL tumour antigen and immunogenic peptides derived from it

CC are useful for cancer immunotherapy. They have the potential to induce an

CC immune response, by eliciting a CTL response. The DNA molecule is used to

CC construct recombinant or fusion proteins.

XX

SQ Sequence 9 AA;

Query Match 91.1%; Score 41; DB 21; Length 9;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MLAAQERRV 10

Db 1 MLAAQERRV 9

|||||

RESULT 3

AAAY05982

ID AAY05982 standard; Peptide; 10 AA.

XX

AC AAY05982;

XX

DT 16-AUG-1999 (first entry)

DE

DE Human cancer antigen NY ESO-1/CAG-3 ORF2 cancer peptide.

XX

XX NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;

KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;

KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;

KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;

KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;

KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;

KW vaccine.

XX

OS Homo sapiens.

XX

PN WO9918206-A2.

XX

PD 15-APR-1999.

XX

PF 21-SEP-1998; 98WO-US19609.

XX

PR 08-OCT-1997; 97US-0061428.

XX

PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX

PI Rosenberg SA, Wang RF;

XX

DR WPI; 1999-277270/23.

DR N-PSDB; AAX58601.

XX

PT Cancer antigen NY ESO1/CAG-3

XX

PS Claim 26; Page 65; 88pp; English.

XX

XX The present sequence represents a cancer peptide that is based on

CC amino acid residues 19-27 of human ESO-1/CAG-3 (or CAG-3) ORF2

CC (see AAY05966), a new and potent tumour antigen capable of eliciting

CC an antigen specific immune response by T cells. Cancer peptides

CC derived from CAG-3 ORF2, CAG-3 ORF1 (see AAY05965), portions of them

CC and their variants (see AAY05967-87), are useful as cancer vaccines

CC that protect against cancer. The invention provides: vectors and

CC host cells (also useful as vaccines); a method of diagnosis of

CC cancer or precancer; a transgenic animal; antisense oligonucleotides

CC that inhibit expression of the cancer peptide or tumour antigen;

CC antibodies reacting with a CAG-3 cancer peptide, useful in

CC diagnostic and detection assays; and methods for preventing or

CC inhibiting cancer by administering a cancer peptide, with or without



CC an HLA molecule. The cancer peptides form part of, or are derived  
 CC from, cancers such as primary or metastatic melanoma, thymoma,  
 CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine  
 CC cancer, cervical cancer, bladder cancer, kidney cancer and  
 CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and  
 CC thyroid cancers. Melanoma is treated by inducing cancer-specific T  
 CC cells in vitro for subsequent return to a patient.  
 XX  
 SQ Sequence 10 AA;

Query Match 80.0%; Score 36; DB 20; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 1.4;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LAAQERRV 10  
 |||||  
 Db 1 LAAQERRV 8

RESULT 4  
 AAB31332  
 ID AAB31332 standard; peptide; 10 AA.  
 XX  
 AC AAB31332;  
 XX  
 DT 20-APR-2001 (first entry)  
 XX  
 OS Exemplary antigen characteristic of tumours and derived from NY-ESO-1.  
 DE  
 XX  
 KW MAGE-A1; HLA; human leukocyte antigen; CD4+ T lymphocyte; cancer;  
 KW MAGE-A1 HLA class II-binding protein; vaccine.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200078806-A1.  
 XX  
 PD 28-DEC-2000.  
 XX  
 PF 14-JUN-2000; 2000WO-US16287.  
 XX  
 PR 18-JUN-1999; 99US-0336091.  
 XX  
 PA (LUDW-) LUDWIG INST CANCER RES.  
 XX  
 PI Van Snick J, Lethe B, Chaux P, Boon-Falleur T, Van Der Bruggen P;  
 XX  
 DR WPI; 2001-102698/11.  
 XX

Novel MAGE-A1 human leukocyte antigen class II peptides which bind to  
 and are presented to the class II molecules, useful for inducing immune  
 response and treating cancers characterized by expression of MAGE-A1 -  
 PS Disclosure; Page 32; 78pp; English.

XX AAB31302-59 represent exemplary antigens which are characteristic  
 CC of tumours. They can be used to enhance the immune response of vaccines  
 CC comprising peptides derived from human MAGE-A1 HLA (human leukocyte  
 CC antigen) class II-binding protein. Peptides derived from the MAGE-A1 HLA  
 CC binding protein stimulate the activity and proliferation of CD4+ T  
 CC lymphocytes. The MAGE-A1 HLA binding protein is useful as a diagnostic  
 CC agent for diagnosing a disorder characterized by expression of MAGE-A1.  
 CC The protein is used for treating a disorder characterized by expression  
 CC of MAGE-A1 such as cancers e.g. melanoma, squamous cell carcinomas,  
 CC colorectal carcinomas, osteosarcomas, and lymphocytic leukemias.  
 CC Peptides derived from the MAGE-A1 HLA binding protein are useful in the  
 CC production of anti-tumour vaccines.  
 XX  
 SQ Sequence 10 AA;

Query Match 80.0%; Score 36; DB 22; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 1.4;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LAAQERRV 10  
 |||||  
 Db 1 LAAQERRV 8

RESULT 5  
 ABG66803  
 ID ABG66803 standard; Peptide; 10 AA.  
 XX  
 AC ABG66803;  
 XX  
 DT 24-SEP-2002 (first entry)  
 XX  
 DE Tumour antigen NY-ESO-1/CAG3 ORF2, HLA-A31 epitope.  
 XX  
 KW Beta-2 microglobulin; beta-2m; cytotoxic T lymphocyte; CTL; HLA;  
 KW human leukocyte antigen; fusion protein; epitope; cytostatic; tumour;  
 KW gastrointestinal tumour; colorectal cancer; gastro-oesophageal cancer;  
 KW liver cancer; biliary tract cancer; pancreatic cancer; vaccine;  
 KW prostatic cancer; testicular cancer; lung cancer; breast cancer;  
 KW malignant melanoma; mesothelioma; brain tumour; ovarian cancer;  
 KW uterine cancer; cervical cancer; head and neck cancer; bladder cancer;  
 KW Kaposi's sarcoma; renal carcinoma; leukaemia; lymphoma;  
 KW acquired immunodeficiency syndrome; AIDS-related lymphoma.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200236146-A2.  
 XX  
 PD 10-MAY-2002.  
 XX  
 PF 01-NOV-2001; 2001WO-GB04844.  
 XX  
 PR 02-NOV-2000; 2000GB-0026812.  
 XX  
 PA (ISIS-) ISIS INNOVATION LTD.  
 XX  
 PI Tafuro S, Meier U, McMichael AJ, Bell JI, Layton G, Hunter M;  
 XX  
 DR WPI; 2002-508108/54.  
 XX

New polynucleotide capable of expressing an epitope-beta2m fusion  
 protein useful for generating cytotoxic T lymphocyte responses against  
 a tumour and in restoring antigen presentation in the tumour of a host  
 -  
 PS Disclosure; Page 25; 46pp; English.

XX The invention relates to a new polynucleotide capable of expressing an  
 CC epitope-beta-2m fusion protein useful for generating cytotoxic T  
 CC lymphocyte (CTL) responses against a tumour or in restoring antigen  
 CC presentation in the tumour of a host. Also included are a polynucleotide  
 CC capable of expressing an epitope-beta-2m fusion protein in combination  
 CC with a vaccination agent that stimulates a CTL response against the  
 CC epitope of the fusion protein for simultaneous, separate or sequential  
 CC use in the treatment of cancer and a method of treating a tumour by  
 CC administering a capable of expressing an epitope-beta-2m fusion protein,  
 CC and optionally a vaccination agent that stimulates a CTL response against  
 CC the epitope of the fusion protein. The polynucleotide is useful for  
 CC generating CTL responses against tumours, for restoring antigen  
 CC presentation in the tumour, and subsequently for treating cancers, such  
 CC as gastrointestinal tumour, prostatic, testicular, lung or breast cancer,  
 CC malignant melanoma, mesothelioma, brain tumour, ovarian cancer, uterine  
 CC cancer including cervical cancer, cancer of the head and neck, bladder  
 CC cancer, Kaposi's sarcoma, AIDS (acquired immunodeficiency syndrome)-  
 CC related Kaposi's sarcoma, sarcomas, osteosarcoma, renal carcinoma, and  
 CC haematopoietic malignant tumours such as leukaemia and lymphoma.  
 CC The epitope is an HLA (human leukocyte antigen) peptide derived from a  
 CC viral or tumour antigen. The present sequence is a tumour HLA epitope  
 CC used in the fusion proteins of the invention.  
 XX  
 SQ Sequence 10 AA;

Query Match 80.0%; Score 36; DB 23; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 1.4;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAAQERRV 10  
 Db 1 LAAQERRV 8  
 |||||

RESULT 6  
 AAY05981  
 ID AAY05981 standard; Peptide; 9 AA.  
 XX AC AAY05981;  
 XX DT 16-AUG-1999 (first entry)  
 XX DE Human cancer antigen NY ESO-1/CAG-3 ORF2 cancer peptide.  
 XX KW NY ESO-1/CAG-3 gene; cancer peptide; antigen; human;  
 KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
 KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
 KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
 KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
 KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
 KW vaccine.  
 XX OS Homo sapiens.  
 XX PN W09918206-A2.  
 XX PD 15-APR-1999.  
 XX PF 21-SEP-1998; 98WO-US19609.  
 XX PR 08-OCT-1997; 97US-0061428.  
 XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX PI Rosenberg SA, Wang RF;  
 XX DR WPI: 1999-277270/23.  
 XX DR N-PSDB; AAX58602.  
 XX PT Cancer antigen NY ESO1/CAG-3  
 XX PS Claim 27; Page 65; 88pp; English.  
 CC The present sequence represents a cancer peptide that corresponds  
 CC to amino acid residues 19-27 of human ESO-1/CAG-3 (or CAG-3) ORF2  
 CC (see AAY05986), a new and potent tumour antigen capable of eliciting  
 CC an antigen specific immune response by T cells. Cancer peptides  
 CC derived from CAG-3 ORF2, CAG-3 ORF1 (see AAY05965), portions of them  
 CC and their variants (see AAY05967-87), are useful as cancer vaccines  
 CC that protect against cancer. The invention provides: vectors and  
 CC host cells (also useful as vaccines); a method of diagnosis of  
 CC cancer or precancer; a transgenic animal; antisense oligonucleotides  
 CC that inhibit expression of the cancer peptide or tumour antigen;  
 CC antibodies reacting with a CAG-3 cancer peptide, useful in  
 CC diagnostic and detection assays; and methods for preventing or  
 CC inhibiting cancer by administering a cancer peptide, with or without  
 CC an HLA molecule. The cancer peptides form part of, or are derived  
 CC from, cancers such as primary or metastatic melanoma, thymoma,  
 CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine  
 CC cancer, cervical cancer, bladder cancer, kidney cancer and  
 CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and  
 CC thyroid cancers. Melanoma is treated by inducing cancer-specific T  
 CC cells in vitro for subsequent return to a patient.  
 XX SQ Sequence 9 AA;

Query Match 71.1%; Score 32; DB 20; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AAQERRV 10  
 Db 1 AAQERRV 7  
 |||||

RESULT 7  
 AAO15266  
 ID AAO15266 standard; peptide; 10 AA.  
 XX AC AAO15266;  
 XX DT 16-SEP-2002 (first entry)  
 XX DE Immune system modulating peptide 1.  
 XX KW Immune system modulation; infection; atopic state; leukocytic disorder;  
 KW vaccination augmentation; anaemia; AIDS; HIV; cytomegalovirus;  
 KW herpes simplex virus; Epstein-Barr virus; HTLV; Marek's disease;  
 KW hog cholera virus; feline sarcoma virus; distemper virus;  
 KW Candida albicans; Aspergillosis.  
 XX OS Synthetic.  
 XX PN US2002052321-A1.  
 XX PD 02-MAY-2002.  
 XX PF 31-MAY-2001; 2001US-0872503.  
 XX PR 07-JUN-1995; 95US-0484511.  
 XX PR 04-AUG-1999; 99US-0368449.  
 XX PR 28-OCT-1991; 91US-0783517.  
 XX PR 02-JAN-1992; 92US-0816205.  
 XX PR 28-OCT-1992; 92US-0967633.  
 XX PR 28-OCT-1993; 93US-0144779.  
 XX PA (GREE/) GREEN L R.  
 PA (SINA/) SINACKEVICH N V.  
 PA (IVAN/) IVANOV V T.  
 PA (MIKH/) MIKHAYLOVA I I.  
 PA (VASKH/) VASKOVSKY B V.  
 PA (MIKH/) MIKHALTSOV A N.  
 PA (KHAV/) KHAVINSON V K.  
 PA (MORO/) MOROZOV V G.  
 XX PI Green LR, Sinackevich NV, Ivanov VT, Mikhalyova II, Vaskovsky BV;  
 PI Mikhaltsov AN, Khavinson VK, Morozov VG;  
 XX DR WPI: 2002-499329/53.  
 XX PT New peptide-containing pharmaceutical compositions exhibit a broad  
 XX range of efficacy for modulation of the immune system, useful in the  
 XX treatment of infection  
 XX Disclosure: Page 2; 26pp; English.  
 CC The invention comprises compositions containing a peptide of the formula  
 CC R'-Glx-Glx-Lys-R'', the compositions of the invention are useful for  
 CC modulating the activity of a host's immune system, to treat infections,  
 CC atopic states, leukocytic disorders and for augmenting vaccinations in a  
 CC host. The compositions of the invention are useful for the treatment of  
 CC anaemia, AIDS, HIV, cytomegalovirus, herpes simplex virus; Epstein-Barr  
 CC virus, HTLV, Marek's disease, hog cholera virus, feline sarcoma virus,  
 CC distemper virus, Candida albicans and Aspergillosis. The present amino  
 CC acid sequence represents a peptide of the invention.  
 XX SQ Sequence 10 AA;

QY 1 AMLAAQERR 9  
 ||| |:::  
 Db 1 AMLTAZZK 9

## RESULT 8

AAAB09858  
 ID AAB09858 standard; peptide; 9 AA.

AC AAB09858;

DT 12-OCT-2000 (first entry)

DE Immunomodulatory peptide SEQ ID 27.

KW Immunomodulatory peptide; AIDS; acquired immunodeficiency syndrome;  
 infection; leukocytic disorder.

XX Synthetic.

OS US0606622-A.

PN 23-MAY-2000.

PD 28-OCT-1993; 93US-0144779.

PF 28-OCT-1991; 91US-0783517.

PR 02-JAN-1992; 92US-0816205.

PR 28-OCT-1992; 92US-0967633.

XX (CYTR-) CYTRAN INC.

XX Green LR, Ivanov VT, Mikhalyova II, Vaskovsky BV, Mikhailov AN;

PI Khavinson VK, Morozov VG, Sinackevich NV;

XX WPI; 2000-410680/35.

DR Modulating immune system activity, useful for restoring the natural  
 balance of the system and for treating diseases associated with  
 suppressed or hyperactive immune systems, by administering a peptide  
 with at least five amino acids -

PS Disclosure; Column 37; 35pp; English.

XX The present sequence is a peptide which can be used as an  
 immunomodulator. This can be used to treat primary immunodeficiencies  
 such as AIDS, DeGeorge's syndrome and severe combined immunodeficiency,  
 secondary immunodeficiencies including energy from tuberculosis,  
 drug-induced leukopenia, non-HIV viral illnesses leukopenia, radiation  
 poisoning, toxin exposure and malnutrition, hyperactive immune states  
 including systemic lupus erythematosus, rheumatic fever, rheumatoid  
 arthritis and multiple sclerosis, bacterial, viral, fungal and parasitic  
 infections, and leukocytic disorders such as pre-leukaemias and leukemoid  
 reactions. Also, some anaemias may be treated, including acute  
 haemorrhagic anaemia, anaemias of chronic disease, megaloblastic  
 anaemias, iron deficiency anaemias and haemoglobinopathies.

XX Sequence 9 AA;

Query Match 55.6%; Score 25; DB 21; Length 9;  
 Best Local Similarity 44.4%; Pred. No. 9.3e+05;

Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 MAAQERR 10

||| |:::  
 Db 1 MLTAZZKAV 9

## RESULT 9

AAE20695

ID AAE20695 standard; peptide; 9 AA.

XX

AC AAE20695;

01-JUL-2002 (first entry)

DE Immunomodulatory peptide #27 used to treat immune disorders.

XX Therapy; immunodeficient; hyperactive immune state; leprosy; infection;  
 KW spirochetal infection; tuberculosis; human immunodeficiency virus; HIV;  
 KW aspergillosis; syphilis; Marek's disease; schistosomiasis; toxoplasmosis;  
 KW malaria; leishmaniasis; pneumocystis; leukocytic disorder; wound healing;  
 KW immune system; acquired immune deficiency syndrome; malnutrition; AIDS;  
 KW anaemia; tuberculosis; DeGeorge's syndrome; leukopenia; toxin exposure;  
 KW radiation poisoning; multiple sclerosis; rheumatoid arthritis; vulvovaginal;  
 KW cytostatic; dermatological; prophylaxis; systemic lupus erythematosus;  
 KW severe combined immunodeficiency; immunomodulatory peptide.

XX Unidentified.

OS US6346514-B1.

PN 12-FEB-2002.

PD 04-AUG-1999; 99US-0368449.

PF 28-OCT-1993; 93US-0144779.

PR 28-OCT-1991; 91US-0783517.

PR 02-JAN-1992; 92US-0816205.

PR 28-OCT-1992; 92US-0967633.

XX (CYTR-) CYTRAN INC.

XX Green LR, Sinackevich NV, Ivanov VT, Mikhalyova II, Vaskovsky BV;

PI Mikhailov AN, Khavinson VK, Morozov VG;

XX WPI; 2002-266497/31.

PT Treating bacterial, fungal and parasitic infections and immune  
 disorders in a host by administering an immunomodulatory peptide -

PS Disclosure; Column 37; 30pp; English.

XX The present invention relates to compositions and methods for the therapy  
 of immunodeficient, immunodepressed or hyperactive immune states. The  
 method of treating an infection in a host comprises administering an  
 immunomodulatory peptide to the host. The method is useful for treating  
 bacterial infections such as mycobacterial infections (e.g. tuberculosis,  
 leprosy), spirochetal infections (e.g. syphilis), viral infections such  
 as Marek's disease, HIV (human immunodeficiency virus) 1 and 2, fungal  
 infections such as candida albicans and aspergillosis and parasitic  
 diseases such as malaria, schistosomiasis, toxoplasmosis, leishmaniasis  
 and pneumocystis. They are also used to treat anaemias, atopic states,  
 leukocytic disorders (e.g. pre-leukaemias) and to enhance wound healing.  
 The method enhances or suppresses the immune system and immunological  
 enhancement is employed as a treatment for variety of primary diseases  
 (e.g., acquired immune deficiency syndrome (AIDS), DeGeorge's syndrome,  
 severe combined immunodeficiency), secondary immunodeficiencies (e.g.  
 energy from tuberculosis and drug-induced leukopenia), toxin exposure,  
 radiation poisoning, and malnutrition. Immunomodulation by the methods  
 of the invention is also useful for treating hyperactive immune states  
 such as systemic lupus erythematosus, rheumatoid arthritis and multiple  
 sclerosis. The immunomodulatory peptides of the invention can be  
 administered to a patient susceptible or at risk for infection, anaemia  
 or other disorders. They can be used for surgical prophylaxis to lessen  
 the risk of infectious complications and enhance the host's restorative  
 response to blood loss. The present sequence is an immunomodulatory  
 peptide of the invention.

XX Sequence 9 AA;

Query Match 55.6%; Score 25; DB 23; Length 9;  
 Best Local Similarity 44.4%; Pred. No. 9.3e+05;

Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 MAAQERV 10  
 ID ABP75370 standard; Peptide; 9 AA.  
 AC ABP75370;  
 DE 20-FEB-2003 (first entry)  
 XX Chlamydia trachomatis peptide epitope #76.  
 KW Antibacterial; secreted protein; intracellular bacterium.  
 OS Chlamydia trachomatis.  
 PN WO200282091-A2.  
 PD 17-OCT-2002.  
 PF 09-APR-2002; 2002WO-DK00234.  
 PR 09-APR-2001; 2001DK-0000581.  
 XX 09-APR-2001; 2001US-282513P.  
 PA (SHAW/) SHAW A C.  
 XX (VAND/) VANDAHN B B.  
 PI Shaw AC, Vandahl BB;  
 XX WPI; 2003-058585/05.  
 DR Identifying intracellular bacterial proteins by labeling proteins in  
 XX the presence of a eukaryotic protein synthesis inhibitor, performing  
 PT electrophoresis, autoradiography and comparing profiles to an  
 PT infected-cell lysate profile -  
 XX Claim 54; Page 174; 179pp; English.  
 PS The present invention relates to a method (M1) for identifying secreted  
 CC intracellular bacterial proteins (BP). M1 comprises: (a) selectively  
 CC visualising BP by pulse labelling in the presence of an inhibitor of  
 CC eukaryotic protein synthesis followed by 2D electrophoresis and  
 CC autoradiography; (b) comparing protein profiles (PF) of purified bacteria  
 CC to PF of total lysate (TL) of infected cells; and (c) identifying protein  
 CC spots present in differential images from gels loaded with TL. The  
 CC present sequence is one such bacterial peptide epitope which was  
 CC identified by the method of the invention.  
 XX Sequence 9 AA;  
 SQ  
 Query Match 53.3%; Score 24; DB 24; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 9.3e+05;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 3 LAAQERV 10  
 ID LAALENV 9  
 AC AAG62638;  
 DE 11-SEP-2001 (first entry)  
 XX Hsp47 HLA-A2 homologous region peptide.  
 DE

KW Immune system cell: ex vivo expanded anti-angiogenic T cell; EAT cell;  
 KW immunotherapy; cancer; neovascularisation; AV malformation; infection;  
 KW endometriosis; autoimmune disease; psoriasis; rheumatoid arthritis.  
 XX Unidentified.  
 OS  
 PN WO200137671-A1.  
 XX 31-MAY-2001.  
 PD 22-NOV-2000; 2000WO-US32322.  
 XX 24-NOV-1999; 99US-0167513.  
 PF (HOPE/) HOPE E G.  
 XX Hope EG;  
 PI WPI; 2001-367593/38.  
 DR Composition for treating cellular proliferative disorders, including  
 XX early and late stage cancers and solid tissue tumors, comprises ex vivo  
 XX expanded anti-angiogenic T cells -  
 XX Example 11; Page 32; 53pp; English.  
 XX The present invention relates to a composition comprising ex vivo  
 CC expanded anti-angiogenic T cells (EAT cells) which selectively damage  
 CC tumour-associated vasculature. These immune system derived cells can thus  
 CC be used in the treatment of many types of cancer, other diseases  
 CC associated with neovascularisation, such as AV malformation, and  
 CC endometriosis, autoimmune diseases such as rheumatoid arthritis and  
 CC psoriasis, and infections causing neovascularisation, including  
 CC echinococcus granulosis and sarcoid. The present sequence is a peptide  
 CC fragment of Hsp47 homologous to HLA-A2, which was used in the  
 CC exemplification of the invention.  
 XX Sequence 9 AA;  
 SQ  
 Query Match 51.1%; Score 23; DB 22; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 9.3e+05;  
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AMIAQER 8  
 DB 1 AVLSAQER 8  
 RESULT 12  
 AAG97495  
 ID AAG97495 standard; Peptide; 10 AA.  
 XX AAG97495;  
 AC AAG97495;  
 XX 18-SEP-2001 (first entry)  
 DT Human complementary peptide, SEQ ID NO: 3690.  
 DE Human complementary peptide; ligand; drug discovery; drug design.  
 XX Homo sapiens.  
 OS WO200142277-A2.  
 PN 14-JUN-2001.  
 PD 13-DEC-2000; 2000WO-GB04776.  
 XX 13-DEC-1999; 99GB-0029464.  
 PF (PROT-) PROTEOM LTD.  
 XX Roberts GW, Heal JR;  
 PI

XX WPI; 2001-408419/43.

XX A set of peptide ligands consisting of specific complementary peptides

PT to proteins encoded by genes of the human genome, useful in an assay

PT for screening and identifying of one or more novel peptides which are

PT drug candidates or pro-drugs -

XX Example 6; Page 572; 646pp; English.

PS The invention relates to a set of complementary peptide ligands

XX generated from the human genome. The complementary peptides

CC interact with their relevant target proteins encoded in the human

CC genome. They can be used as reagents in drug discovery and as lead

CC ligands to facilitate drug design and development. The present

XX sequence is a complementary peptide provided in the specification.

XX Sequence 10 AA;

Query Match 51.1%; Score 23; DB 22; Length 10;

Best Local Similarity 57.1%; Pred. No. 4.6e+02;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 AAQERRV 10

DB 1 AEEERRI 7

RESULT 13

AAAB07254

ID AAB07254 standard; peptide; 7 AA.

XX AAB07254;

AC 12-OCT-2000 (first entry)

DT Immunomodulatory peptide SEQ ID 24.

XX Immunomodulatory peptide; AIDS; acquired immunodeficiency syndrome;

KW infection; leukocytic disorder.

KW Synthetic.

OS US6066622-A.

PN 23-MAY-2000.

PD 28-OCT-1993; 93US-0144779.

PF 28-OCT-1991; 91US-0783517.

XX 02-JAN-1992; 92US-0816205.

PR 28-OCT-1992; 92US-0967633.

XX (CYTR-) CYTRAN INC.

PA Green LR, Ivanov VT, Mikhalyova II, Vaskovsky BV, Mikhaltsov AN;

PI Khavinson VK, Morozov VG, Sinackevich NV;

XX WPI; 2000-410680/35.

Modulating immune system activity, useful for restoring the natural

PT balance of the system and for treating diseases associated with

PT suppressed or hyperactive immune systems, by administering a peptide

PT with at least five amino acids -

XX Disclosure; Column 37; 35pp; English.

XX The present sequence is a peptide which can be used as an

CC immunomodulator. This can be used to treat primary immunodeficiencies

CC such as AIDS, DeGeorge's syndrome and severe combined immunodeficiency,

CC secondary immunodeficiencies including anergy from tuberculosis,

CC drug-induced leukopenia, non-HIV viral illnesses leukopenia, radiation

CC poisoning, toxin exposure and malnutrition, hyperactive immune states

CC including systemic lupus erythematosus, rheumatic fever, rheumatoid

CC arthritis and multiple sclerosis, bacterial, viral, fungal and parasitic

CC infections, and leukocytic disorders such as pre-leukaemias and leukemoid

CC reactions. Also, some anaemias may be treated, including acute

CC haemorrhagic anaemia, anaemias of chronic disease, megaloblastic

CC anaemias, iron deficiency anaemias and haemoglobinopathies.

XX Sequence 7 AA;

Query Match 48.9%; Score 22; DB 21; Length 7;

Best Local Similarity 42.9%; Pred. No. 9.3e+05;

Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 MAAQER 8

DB 1 MLTAZK 7

RESULT 14

AAAE20692

ID AAE20692 standard; peptide; 7 AA.

XX AAE20692;

AC 01-JUL-2002 (first entry)

DT Immunomodulatory peptide #24 used to treat immune disorders.

XX Therapy; immunodeficient; hyperactive immune state; leprosy; infection;

KW spirochetal infection; tuberculosis; human immunodeficiency virus; HIV;

KW aspergillosis; syphilis; Marek's disease; schistosomiasis; toxoplasmosis;

KW malaria; leishmaniasis; pneumocystis; leukocytic disorder; wound healing;

KW immune system; acquired immune deficiency syndrome; malnutrition; AIDS;

KW anaemia; tuberculosis; DeGeorge's syndrome; leukopenia; toxin exposure;

KW radiation poisoning; multiple sclerosis; rheumatoid arthritis; vulnerability;

KW cytostatic; dermatological; prophylaxis; systemic lupus erythematosus;

XX severe combined immunodeficiency; immunomodulatory peptide.

OS Unidentified.

XX US6346514-B1.

PN 12-FEB-2002.

PD 04-AUG-1999; 99US-0368449.

PF 28-OCT-1993; 93US-0144779.

XX 28-OCT-1991; 91US-0783517.

PR 02-JAN-1992; 92US-0816205.

XX 28-OCT-1992; 92US-0967633.

XX (CYTR-) CYTRAN INC.

PA Green LR, Sinackevich NV, Ivanov VT, Mikhalyova II, Vaskovsky BV;

PI Mikhaltsov AN, Khavinson VK, Morozov VG;

XX WPI; 2002-266497/31.

Treating bacterial, fungal and parasitic infections and immune

PT disorders in a host by administering an immunomodulatory peptide -

PS Disclosure; Column 37; 30pp; English.

XX The present invention relates to compositions and methods for the therapy

CC of immunodeficient, immunodepressed or hyperactive immune states. The

CC method of treating an infection in a host comprises administering an

CC immunomodulatory peptide to the host. The method is useful for treating

CC bacterial infections such as mycobacterial infections (e.g. tuberculosis,

CC leprosy), spirochetal infections (e.g. syphilis), viral infections such

CC as Marek's disease, HIV (human immunodeficiency virus) 1 and 2, fungal

CC infections such as Candida albicans and aspergillosis and parasitic

CC diseases such as malaria, schistosomiasis, toxoplasmosis, leishmaniasis

CC and pneumocystis. They are also used to treat anaemias, atopic states,

CC leukocytic disorders (e.g. pre-leukaemias) and to enhance wound healing.  
 CC The method enhances or suppresses the immune system and immunological  
 CC enhancement is employed as a treatment for variety of primary diseases  
 CC (e.g., acquired immune deficiency syndrome (AIDS), DeGeorge's syndrome,  
 CC severe combined immunodeficiency), secondary immunodeficiencies (e.g.  
 CC anergy from tuberculosis and drug-induced leukopenia), toxin exposure,  
 CC radiation poisoning, and malnutrition. Immunomodulation by the methods  
 CC of the invention is also useful for treating hyperactive immune states  
 CC such as systemic lupus erythematosus, rheumatoid arthritis and multiple  
 CC sclerosis. The immunomodulatory peptides of the invention can be  
 CC administered to a patient susceptible or at risk for infection, anaemia  
 CC or other disorders. They can be used for surgical prophylaxis to lessen  
 CC the risk of infectious complications and enhance the host's restorative  
 CC response to blood loss. The present sequence is an immunomodulatory  
 CC peptide of the invention.

SQ Sequence 7 AA;

Query Match 48.9%; Score 22; DB 23; Length 7;  
 Best Local Similarity 42.9%; Pred. No. 9.3e+05;  
 Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 MLAAQER 8  
 II I:::  
 Db 1 MLTAZZK 7

# RESULT 15

AAR36590  
 ID AAR36590 standard; peptide; 8 AA.

XX AAR36590;

DT 25-MAR-2003 (updated)

DT 26-AUG-1993 (first entry)

DE Immunodepressed state treatment peptide #1.

XX Peptide; treatment; immunodepression; opportunistic infection; AIDS;  
 KW acquired immunodeficiency syndrome; pharmaceutical; salt; influenza;  
 KW restoration; stimulation; immune function; veterinary practice; anemia;  
 KW immunomodulatory; prophylaxis; hypotrophy; farming; animals; burns;  
 KW fur bearing; poultry; respiratory disease; wounds; open sores; rashes;  
 KW allergy; sun exposure; trauma; eczema; psoriasis; bone fractures;  
 KW lesions; gingival disease; gynecological; infection; infaralymphatic;  
 KW iron deficient.

XX Synthetic.

XX WO9308816-A1.

XX 13-MAY-1993.

XX 28-OCT-1992; 92WO-US09252.

XX 28-OCT-1991; 91US-0783517.

XX (CYTO-) CYTOVEN.

XX Green LR;

XX WPI; 1993-167379/20.

XX Treatment of opportunistic infection in immuno-depressed patients -  
 PT by administering simple peptide(s)

XX Disclosure; Page 3; 34pp; English.

XX The sequences given in AAR36590-98 are simple peptides which can be  
 CC used in the treatment of immunodepressed states and of opportunistic  
 CC infections in immunodepressed states associated with acquired immuno-  
 CC deficiency syndrome. The peptides are all covered by the generic  
 CC sequence;

CC where R' = Thr-Ala, Thr-Pro, Ser-Ala, Ser-Pro, Ser-Ser, Met-Leu-Thr-  
 CC Ala, or Leu-Thr-Ala;  
 CC R = -H, -Ala, -Ala-Ala or Ala-Val.

CC The peptides may be prepared by conventional methods of enzyme  
 CC synthesis and may be used as the active ingredients in a  
 CC pharmaceutical preparation as free peptides or in the form of a  
 CC water soluble salt, such as sodium, potassium, ammonium or zinc salt.  
 CC These peptides may be used for the restoration and stimulation of  
 CC immune functions. They may also be used in veterinary practice as an  
 CC immunomodulatory agent for the prophylaxis and treatment of hypo-  
 CC trophy in farming animals, fur bearing animals and poultry. The  
 CC opportunistic infections which may be treated using these peptides  
 CC include respiratory disease, influenza, AIDS, burns, wounds, other  
 CC open sores, rashes due to allergic reaction, sun exposure, local  
 CC trauma, eczema, psoriasis, etc. The peptides may be used to assist  
 CC healing in immunodepressed states such as healing bone fractures,  
 CC lesions, gingival diseases, gynecological infections, infaralymphatic  
 CC infections, etc. These peptides are particularly useful in the  
 CC treatment of anemias of all types such as those caused by loss of  
 CC blood, blood formation affection or blood destruction enhancement  
 CC including all types of iron deficient anemia.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 CC (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 8 AA;

Query Match 48.9%; Score 22; DB 14; Length 8;  
 Best Local Similarity 42.9%; Pred. No. 9.3e+05;  
 Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 MLAAQER 8  
 II I:::  
 Db 1 MLTAZZK 7

Search completed: October 7, 2003, 14:27:45  
 Job time : 58.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:24:35 ; Search time 24 Seconds  
(without alignments)  
40.070 Million cell updates/sec

Title: US-09-807-512-25  
Perfect score: 45  
Sequence: 1 AMLAAQRRRV 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues  
Total number of hits satisfying chosen parameters: 1100

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_76.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	DB ID	Description
1	20	44.4	10	4 S14943	UGA3 leader peptid
2	17	37.8	8	4 I55411	hypothetical hist
3	15	33.3	6	2 A49792	acylaminoacyl-pept
4	15	33.3	8	2 A54823	olfactory receptor
5	15	33.3	8	2 B54823	olfactory receptor
6	15	33.3	8	2 I64832	Ca2+-transporting
7	15	33.3	10	2 S70721	heat shock protein
8	15	33.3	10	2 D33098	214K exoantigen (v
9	15	33.3	10	2 PC4374	telomeric and tetr
10	14	31.1	7	2 PH0932	T-cell receptor be
11	14	31.1	9	2 A44873	caldesmon - rabbit
12	14	31.1	10	2 PH0113	alpha-amylase (EC
13	14	31.1	10	2 I52645	gene B-50 protein
14	13	28.9	6	2 A19780	transferrin - bovi
15	13	28.9	8	2 S70727	ipgf protein - Shi
16	13	28.9	8	2 I48935	apolipoprotein A-I
17	13	28.9	9	2 S02384	probable membrane
18	13	28.9	9	2 S66608	quinoline 2-oxid
19	13	28.9	10	2 S09387	PVI protein - huma
20	13	28.9	10	2 S66458	ferritin - Rhizo
21	13	28.9	10	2 E49033	T-cell receptor ga
22	13	28.9	10	2 PH0933	T-cell receptor be
23	13	28.9	10	2 P00785	NADH2 dehydrogen
24	12	26.7	5	2 S55237	zinc-binding prote
25	12	26.7	5	2 PT0672	T-cell receptor be
26	12	26.7	6	2 PC4127	hypothetical 6 pro
27	12	26.7	6	2 PT0510	T-cell receptor be
28	12	26.7	7	2 P50254	18K protein 5507 -
29	12	26.7	8	2 PT0323	Ig heavy chain CRD

30	12	26.7	9	2 A43848	cell surface adhes
31	12	26.7	9	2 D57444	neuropeptide Grb-A
32	12	26.7	9	2 A45199	L-lysophorin - Ja
33	12	26.7	9	2 C60070	gastrin - domestic
34	12	26.7	9	2 A33527	fructose-2,6-bisph
35	12	26.7	9	4 S15594	orf 1 rara 5'-regi
36	12	26.7	10	2 S24190	tryptase (EC 3.4.2
37	12	26.7	10	2 PNO136	pepsin (EC 3.4.23
38	12	26.7	10	2 GXHU1	gastric juice pept
39	12	26.7	10	2 D37397	hypothetical prote
40	12	26.7	10	2 PS0451	24K protein 4302 -
41	12	26.7	10	2 PH0807	T-cell receptor al
42	12	26.7	10	2 PH0894	myosin light chain
43	11	24.4	5	2 I50385	RNA-directed DNA p
44	11	24.4	6	2 A35890	alpha-tubulin - Ch
45	11	24.4	6	2 I48126	

ALIGNMENTS

RESULT 1

S14943  
UGA3 leader peptide - yeast (Saccharomyces cerevisiae)  
C:Species: Saccharomyces cerevisiae  
C:Date: 07-Aug-1998 #sequence\_revision 07-Aug-1998 #text\_change 08-Dec-2000  
C:Accession: S14943; A59281  
R.Andre, B.

Mol. Gen. Genet. 220, 269-276, 1990

A>Title: The UGA3 gene regulating the GABA catabolic pathway in Saccharomyces cerevis

A:Reference number: S14943; MUID:90220499; PMID:2109179

A:Accession: S14943

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-10 <AND1>

A:Cross-references: EMBL:X51664; NID:g4747

A:Accession: A59281

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 'HVMMPKQSF' <AND2>

A:Cross-references: EMBL:X51664; NID:g4747; PIDN:CAA35975.1; PID:g1326024

A>Note: erroneous translation for authors' feature in GenBank entry SCUGA3G, release

Query Match 44.4%; Score 20; DB 4; Length 10;  
Best Local Similarity 44.4%; Pred No. 5e+02;  
Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 MLAAQRRRV 10  
| | | | |  
Db 1 MYGQENKV 9

RESULT 2

I55411

hypothetical histone H2A.X (mistranslated) - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 28-Jan-2000 #sequence\_revision 28-Jan-2000 #text\_change 28-Jan-2000

C:Accession: I55411

R.Ivanova, V.S.; Hatch, C.L.; Bonner, W.M.

J. Biol. Chem. 269, 24189-24194, 1994

A>Title: Characterization of the human histone H2A.X gene. Comparison of its promoter

A:Reference number: I55411; MUID:95014156; PMID:7929075

A:Accession: I55411

A>Status: translation not shown; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-8 <IVA>

A:Cross-references: GB:S73863; NID:g765295; PIDN:AAAD14141.1; PID:g4261841

A>Note: this is a hypothetical translation of a sequence from the promoter region pre

C:Genetics:

A:Gene: H2A.X

Query Match 37.8%; Score 17; DB 4; Length 8;  
Best Local Similarity 66.7%; Pred. No. 2.8e+05;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 LAAQER 8  
|||:|  
Db 3 LAARAR 8

## RESULT 3

A49792  
C:Title: N-terminal sequence analysis of N(alpha)-acetylated proteins after unblocking  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 07-May-1999  
C:Accession: A49792  
R:Krishna, R.G.; Chin, C.C.Q.; Wold, F.  
Anal. Biochem. 199, 45-50, 1991  
A:Title: N-terminal sequence analysis of N(alpha)-acetylated proteins after unblocking  
A:Reference number: A49792; MUID:92222120; PMID:1807161  
A:Accession: A49792  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-6 <KRI>  
C:Keywords: acetylated amino end; hydrolase; omega peptidase  
F:1/Modified site: acetylated amino end (Met) #status experimental

Query Match 33.3%; Score 15; DB 2; Length 6;  
Best Local Similarity 75.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 ERVV 10  
|||:|  
Db 2 ERQV 5

## RESULT 4

A54823  
C:Title: olfactory receptor I7 - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 28-Apr-1995 #sequence\_revision 28-Apr-1995 #text\_change 17-Mar-1999  
C:Accession: A54823  
R:Chess, A.; Simon, I.; Cedar, H.; Axel, R.  
Cell 78, 823-834, 1994  
A:Title: Allelic inactivation regulates olfactory receptor gene expression.  
A:Reference number: A54823; MUID:94373818; PMID:8087849  
A:Accession: A54823  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-8 <CHE>

Query Match 33.3%; Score 15; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 ERR 9  
|||  
Db 2 ERR 4

## RESULT 5

B54823  
C:Title: olfactory receptor I7 - western wild mouse (fragment)  
C:Species: Mus spretus (western wild mouse)  
C:Date: 28-Apr-1995 #sequence\_revision 28-Apr-1995 #text\_change 17-Mar-1999  
C:Accession: B54823  
R:Chess, A.; Simon, I.; Cedar, H.; Axel, R.  
Cell 78, 823-834, 1994  
A:Title: Allelic inactivation regulates olfactory receptor gene expression.  
A:Reference number: A54823; MUID:94373818; PMID:8087849  
A:Accession: B54823  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-8 <CHE>

Query Match 33.3%; Score 15; DB 2; Length 8;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 ERR 9  
|||  
Db 2 ERR 4

## RESULT 6

I64832  
C:Title: Ca2+-transporting ATPase (EC 3.6.3.8) - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 19-Apr-2002  
C:Accession: I64832  
R:Wu, K.  
Am. J. Physiol. 264, 333-341, 1993  
A:Title: Molecular cloning and quantification of mRNA and protein encoding Ca2+-ATPase  
A:Reference number: I51892  
A:Accession: I64832  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-8 <RES>  
A:Cross-references: GB:M99223; NID:g203644; PIDN:AAA0992.1; PID:g203646  
C:Genetics:  
A:Gene: SERCALB  
C:Keywords: hydrolase

Query Match 33.3%; Score 15; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 ERR 9  
|||  
Db 5 ERR 7

## RESULT 7

S70721  
C:Title: heat shock protein C62.5 homolog - Salmonella typhimurium (fragment)  
C:Alternate names: high temperature protein G  
C:Species: Salmonella typhimurium  
C:Date: 11-Mar-1998 #sequence\_revision 17-Apr-1998 #text\_change 17-Apr-1998  
C:Accession: S70721  
R:Qi, S.Y.; Li, Y.; Szyroki, A.; Gilles, I.G.; Moir, A.; O'Connor, C.D.  
Mol. Microbiol. 17, 523-531, 1995  
A:Title: Salmonella typhimurium responses to a bactericidal protein from human neutro  
A:Reference number: S70719; MUID:96100451; PMID:8559071  
A:Accession: S70721  
A:Molecule type: protein  
A:Residues: 1-10 <QIS>  
A:Experimental source: strain SL1344  
C:Keywords: ATP binding; heat shock; molecular chaperone

Query Match 33.3%; Score 15; DB 2; Length 10;  
Best Local Similarity 42.9%; Pred. No. 5.6e+03;  
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 LAAQERR 9  
:|||  
Db 1 MKGQETR 7

## RESULT 8

D33098  
C:Title: 21k exoantigen (version 1) - malaria parasite (Plasmodium falciparum) (fragments)  
C:Species: Plasmodium falciparum  
C:Date: 24-Aug-1990 #sequence\_revision 24-Aug-1990 #text\_change 09-Jun-2000  
C:Accession: D33098  
R:Nichols, J.H.; Hager, L.P.  
submitted to the Protein Sequence Database, May 1990  
A:Reference number: A33098  
A:Accession: D33098  
A:Status: preliminary



A:Molecule type: protein  
A:Residues: 1-10 <NIC>

Query Match 33.3%; Score 15; DB 2; Length 10;  
Best Local Similarity 66.7%; Pred. No. 5.6e+03;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 MAAQE 7  
Db 1 MLAMLE 6

## RESULT 9

PC4374  
telomeric and tetraplex DNA binding protein qTBP42 IV - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 28-Oct-1997 #sequence\_revision 28-Oct-1997 #text\_change 07-May-1999  
C:Accession: PC4374  
R: Sarig, G.; Weisman-Shomer, P.; Fry, M.  
Biochem. Biophys. Res. Commun. 237, 617-623, 1997  
A:Title: Telomeric and tetraplex DNA binding properties of qTBP42: A homologue of the CA  
A:Reference number: PC4371; MUID:97445086; PMID:9299414  
A:Accession: PC4374  
A:Molecule type: protein  
A:Residues: 1-10 <SAR>  
C:Comment: This protein binds either strand of the telomeric DNA as well as unimolecular

Query Match 33.3%; Score 15; DB 2; Length 10;  
Best Local Similarity 33.3%; Pred. No. 5.6e+03;  
Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 MAAQERV 10  
Db 2 VLDXKEHRL 10

## RESULT 10

PH0932  
T-cell receptor beta chain V-D-J region (clone 3) - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 09-Oct-1992 #sequence\_revision 09-Oct-1992 #text\_change 30-May-1997  
C:Accession: PH0932  
R: Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.  
J. Exp. Med. 174, 1467-1476, 1991  
A:Title: Analysis of T cell receptor beta chains in Lewis rats with experimental allergic  
A:Reference number: PH0891; MUID:92078857; PMID:1836012  
A:Accession: PH0932  
A:Molecule type: mRNA  
A:Residues: 1-7 <GOL>  
A:Experimental source: complete Freund's adjuvant-immunized lymph node  
C:Keywords: T-cell receptor

Query Match 31.1%; Score 14; DB 2; Length 7;  
Best Local Similarity 60.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 AAQER 8  
Db 2 ASPER 6

## RESULT 11

A44873  
caldesmon - rabbit (fragment)  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 31-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 31-Oct-1997  
C:Accession: A44873  
R: Ikebe, M.; Hornick, T.  
Arch. Biochem. Biophys. 288, 538-542, 1991  
A:Title: Determination of the phosphorylation sites of smooth muscle caldesmon by protei  
A:Reference number: A44873; MUID:91378498; PMID:1898046  
A:Accession: A44873  
A>Status: preliminary

A:Molecule type: protein  
A:Residues: 1-9 <IKE>  
A:Experimental source: skeletal myosin  
A:Note: sequence extracted from NCBI backbone (NCBIP:63199)  
C:Superfamily: caldesmon

Query Match 31.1%; Score 14; DB 2; Length 9;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LAAQER 8  
Db 4 LKIER 9

## RESULT 12

PH0113  
alpha-amylase (EC 3.2.1.1) III - rice (fragment)  
C:Species: Oryza sativa (rice)  
C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 07-May-1999  
C:Accession: PH0113  
R: Chiba, Y.; Niede, Y.; Nakajima, T.; Ichishima, E.  
Agric. Biol. Chem. 55, 901-902, 1991  
A:Title: Unique enzymatic properties of alpha-amylase-III from suspension-cultured ri  
A:Reference number: PH0113; MUID:91234351; PMID:1368633  
A:Accession: PH0113  
A:Molecule type: protein  
A:Residues: 1-10 <CHI>  
A:Experimental source: cv. Sasanishiki  
C:Function:  
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds  
A:Pathway: glycogen/starch degradation  
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 31.1%; Score 14; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 9e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 AQE 7  
Db 5 AQE 7

## RESULT 13

IS2645  
gene B-50 protein - rat (fragment)  
C:Species: Rattus sp. (rat)  
C:Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 17-Mar-1999  
C:Accession: IS2645  
R: Eggen, B.J.; Nieleander, H.B.; Rensen-de Leeuw, M.G.; Schotman, P.; Gispén, W.H.; Sc  
Brain Res. Mol. Brain Res. 23, 221-234, 1994  
A:Title: Identification of two promoter regions in the rat B-50/GAP-43 gene.  
A:Reference number: IS2645; MUID:94335354; PMID:8057779  
A:Accession: IS2645  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-10 <RES>  
A:Cross-references: GB:S71492; NID:g560728  
C:Genetics:  
A:Gene: B-50

Query Match 31.1%; Score 14; DB 2; Length 10;  
Best Local Similarity 42.9%; Pred. No. 9e+03;  
Matches 3; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 MAAQER 8  
Db 1 MLCCMR 7

## RESULT 14

A19780  
transferrin - bovine (fragment)

C;Species: Bos primigenius taurus (cattle)  
C;Date: 25-Oct-1987 #sequence\_revision 25-Oct-1987 #text\_change 18-Jun-1993  
C;Accession: A19780  
R;Brock, J.H.; Esparza, I.; Oliver, R.A.; Spooner, R.L.  
Biochem. Genet. 18, 851-860, 1980  
A;Title: Electrophoretic mobility of N- and C-terminal monoferric fragments of bovine trypsin  
A;Reference number: A19780; MUID:81183891; PMID:7225082  
A;Accession: A19780  
A;Molecule type: protein  
A;Residues: 1-6 <BRO>

Query Match 28.9%; Score 13; DB 2; Length 6;  
Best Local Similarity 75.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 7 ERRV 10  
Db 3 ERTV 6

## RESULT 15

S70727  
ipgF protein - Shigella flexneri (fragment)  
C;Species: Shigella flexneri  
C;Date: 15-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 08-Oct-1999  
C;Accession: S70727  
R;Allaoui, A.; Sansonetti, P.J.; Menard, R.; Barzu, S.; Mounier, J.; Phalipon, A.; Parsd  
Mol. Microbiol. 17, 461-470, 1995  
A;Title: MxiG, a membrane protein required for secretion of Shigella spp. Ipa invasins:  
A;Reference number: S70727; MUID:96100445; PMID:8559065  
A;Accession: S70727  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-8 <ALL>  
A;Cross-references: EMBL:248957; NID:g929880; PIDN:CAA8821.1; PID:g929881  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1995  
C;Genetics:  
A;Gene: ipgF

Query Match 28.9%; Score 13; DB 2; Length 8;  
Best Local Similarity 66.7%; Pred. No. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 RRV 10  
Db 6 RRI 8

Search completed: October 7, 2003, 14:31:17  
Job time : 25 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:17:20 ; Search time 13 seconds  
(without alignments)  
36.174 Million cell updates/sec

Title: US-09-807-512-25  
Perfect score: 45  
Sequence: 1 AMLAAQERRV 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 372

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	20	44.4	8	1 UH09_RAT	P56575 rattus norv
2	16	35.6	9	1 ULAK_MOUSE	P99031 mus musculu
3	15	33.3	6	1 ACPL_RABIT	P25154 oryctolagus
4	14	31.1	8	1 B4K_PORGI	P81886 porphyromon
5	14	31.1	8	1 CLP_THICU	P80488 thiobacillu
6	14	31.1	9	1 RS10_SERMA	O68936 serratia ma
7	14	31.1	10	1 MALE_KLEPN	Q05564 klebsiella
8	13	28.9	8	1 UPAL_HUMAN	P30087 homo sapien
9	13	28.9	10	1 QZOG_COMTE	P80466 comamonas t
10	12	26.7	7	1 UH11_RAT	P56576 rattus norv
11	12	26.7	8	1 NS3_MYCTU	P81152 mycobacteri
12	12	26.7	10	1 FIBB_CERSI	P14537 ceratotheri
13	12	26.7	10	1 GAJU_HUMAN	P01358 homo sapien
14	11	24.4	9	1 FARD_CALVO	P41868 calliphora
15	11	24.4	9	1 ULAH_HUMAN	P31934 homo sapien
16	11	24.4	10	1 CATB_SHEEP	P81205 ovis aries
17	11	24.4	10	1 ESL_LACCA	P81758 lactobacill
18	11	24.4	10	1 FARC_CALVO	P41867 calliphora
19	11	24.4	10	1 SPL_HALRO	Q10997 halocynthia
20	10	22.2	7	1 CIA_ENTFA	P11932 enterococcu
21	10	22.2	8	1 NFB_BOVIN	P15507 bos taurus
22	10	22.2	8	1 UC26_MAIZE	P80632 zea mays (m
23	10	22.2	9	1 FARB_PANRE	P82661 panagrellus
24	10	22.2	9	1 FIBB_PAPAN	P19344 papio anubi
25	10	22.2	9	1 FIBB_PAPHA	P19343 papio hamad
26	10	22.2	9	1 FIBB_THEGE	P19342 theropithec
27	10	22.2	9	1 NEUX_HUMAN	P04277 homo sapien
28	10	22.2	10	1 MOSQ_CLYJA	P19962 clypeaster
29	9	20.0	7	1 CARP_MYTED	P10420 mytilus edu
30	9	20.0	7	1 IGAO_DACDE	P06294 dactylium d
31	9	20.0	7	1 LANC_CARUI	P39960 carnobacter
32	9	20.0	8	1 ANG2_BOTJA	Q10582 bothrops ja
33	9	20.0	8	1 CAD1_ENTFA	P13268 enterococcu

## ALIGNMENTS

```

RESULT 1
UH09_RAT          STANDARD;          PRT;          8 AA.
AC P56575;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Unknown protein from 2D-page of heart tissue (Spot P9) (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE.
RC STRAIN=Wistar; TISSUE=Heart;
RA Li X.-P., Pleissner K.-P., Scheler C., Regitz-Zagrosek V., Salikov J.,
RA Jungblut P.R.;
RL Submitted (SEP-1998) to the SWISS-PROT data bank.
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 8.9, ITS MW IS: 42 kDa.
FT NON_TER 8
SQ SEQUENCE 8 AA; 1029 MW; 9E0775A6C4140B06 CRC64;

Query Match          44.4%; Score 20; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 6 QERR 9
Db 1 QERR 4

RESULT 2
ULAK_MOUSE          STANDARD;          PRT;          9 AA.
AC P99031;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Unknown protein from 2D-page of liver tissue (Spot 2D-0014LD)
DE (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RC TISSUE=Liver;
RA Sanchez J.-C., Rouge V., Frutiger S., Hughes G.J., Yan J.X.,
RA Hoogland C., Appel R.D., Binz P.-A., Hochstrasser D.F.,
RA Cowthorne M.;
RL Submitted (AUG-1998) to the SWISS-PROT data bank.
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 6.0, ITS MW IS: 12.5 kDa.
DR SWISS-2DPAGE; P99031; MOUSE.
FT NON_TER 9

```

```

34 9 20.0 8 1 LPMS_STAEP
35 9 20.0 9 1 BS43_SERPL
36 9 20.0 9 1 FAR3_PENMO
37 9 20.0 9 1 HUTU_KLEAE
38 9 20.0 9 1 UN19_CLOPA
39 9 20.0 10 1 AEG1_AGRAE
40 9 20.0 10 1 ANGL_BOTJA
41 9 20.0 10 1 ANGT_BOVIN
42 9 20.0 10 1 ANGT_CHICK
43 9 20.0 10 1 APE_CARGI
44 9 20.0 10 1 COXH_ONCMI
45 9 20.0 10 1 PAPI_PARMA

```

```

P23211 staphylococ
P83375 serratia pl
P83318 penaeus mon
P12381 klebsiella
P81355 clostridium
P83465 agrocyebe ae
Q10581 bothrops ja
P01017 bos taurus
P01018 gallus gall
P80474 capnocytoph
P80331 oncorhynchu
P81863 pardachirus

```

SQ SEQUENCE 9 AA; 1106 MW; E1E842C3240B145A CRC64;

Query Match 35.6%; Score 16; DB 1; Length 9;  
Best Local Similarity 75.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 ERV 10  
| | :  
Db 4 ERV 7

RESULT 3  
ACPH\_RABIT STANDARD; PRT; 6 AA.  
ID ACPH\_RABIT STANDARD; PRT; 6 AA.  
AC P25154;  
DT 01-MAY-1992 (Rel. 22, Created)  
DT 01-MAY-1992 (Rel. 22, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Acylamino-acid-releasing enzyme (EC 3.4.19.1) (Acyl-peptide hydrolase)  
DE (APH) (Acylaminoacyl-peptidase) (Fragment).  
GN APEH.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Muscle;  
RX MEDLINE=92222120; PubMed=1807161;  
RA Krishna R.G., Chin C.C.Q., Wolf F.;  
RT "N-terminal sequence analysis of N alpha-acetylated proteins after  
unblocking with N-acylaminoacyl-peptide hydrolase.";  
RL Anal. Biochem. 199;45:50(1991).  
CC -1- FUNCTION: THIS ENZYME CATALYZES THE HYDROLYSIS OF THE AMINO-  
TERMINAL PEPTIDE BOND OF AN N-ACETYLATED PEPTIDE TO GENERATE  
AN N-ACETYLATED AA AND A PEPTIDE WITH A FREE AMINO-TERMINUS.  
CC IT PREFERENTIALLY CLEAVES OFF AC-ALA, AC-MET AND AC-SER.  
CC -1- CATALYTIC ACTIVITY: Acylaminoacyl-peptide + H(2)O = acylamino acid  
+ peptide.  
CC -1- SUBUNIT: Homotetramer.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S9C.  
DR PIR: A49792; A49792.  
DR MEROPS: S09.004;  
DR InterPro: IPR002471; Prol\_endopep\_ser.  
DR PROSITE: PS00708; PRO-ENDOPEP\_SER; PARTIAL.  
KW Hydrolase; Acetylation.  
FT MOD\_RES 1 1 ACETYLATION.  
FT NON\_TER 6 6  
SQ SEQUENCE 6 AA; 775 MW; 6732D6C40B16F000 CRC64;

Query Match 33.3%; Score 15; DB 1; Length 6;  
Best Local Similarity 75.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 ERV 10  
| | :  
Db 2 ERV 5

RESULT 4  
B44K\_PORGI STANDARD; PRT; 8 AA.  
ID B44K\_PORGI STANDARD; PRT; 8 AA.  
AC P81886;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE 44 kDa immunogenic protein (Fragment).  
OS Porphyromonas gingivalis (Bacteroides gingivalis).  
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;  
OC Porphyromonadaceae; Porphyromonas.  
OX NCBI\_TaxID=837;  
RN [1]  
RP SEQUENCE.

RC STRAIN=VPB 3492;  
RX MEDLINE=20198497; PubMed=10731616;  
RA Norris J.M., Love D.N.;  
RT "Serum antibody responses of cats to soluble whole cell antigens of  
feline Porphyromonas gingivalis";  
RL Vet. Microbiol. 73:37-49(2000).  
CC -1- SIMILARITY: TO P.GINGIVALIS HEMAGGLUTININ A.  
KW Antigen.  
FT NON\_TER 8 8  
SQ SEQUENCE 8 AA; 989 MW; 9554540326CB476D CRC64;

Query Match 31.1%; Score 14; DB 1; Length 8;  
Best Local Similarity 40.0%; Pred. No. 1.3e+05;  
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 6 QERV 10  
| | :  
Db 4 QERNI 8

RESULT 5  
CLP\_THICU STANDARD; PRT; 8 AA.  
ID CLP\_THICU STANDARD; PRT; 8 AA.  
AC P80488;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Chemolithotroph-specific protein (Fragment).  
DE Thiobacillus cuprinus.  
OS Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
OC Comamonadaceae; Thiomonas.  
OX NCBI\_TaxID=36860;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=DSM 5494;  
RA Marin I., Amaro A.M., Jerez C.A., Amils R., Abad J.P.;  
RL Submitted (SEP-1995) to the SWISS-PROT data bank.  
CC -1- MISCELLANEOUS: FOUND SPECIFICALLY IN CELLS CULTURED  
CHEMOLITHOTROPHICALLY.  
FT NON\_TER 8 8  
SQ SEQUENCE 8 AA; 785 MW; 91487B06DDC3D76D CRC64;

Query Match 31.1%; Score 14; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AQE 7  
| | |  
Db 4 AQE 6

RESULT 6  
RS10\_SERMA STANDARD; PRT; 9 AA.  
ID RS10\_SERMA STANDARD; PRT; 9 AA.  
AC O68936;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE 30S ribosomal protein S10 (Fragment).  
GN RPSJ.  
OS Serratia marcescens.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Serratia.  
OX NCBI\_TaxID=615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Noorani S.M., Lindahl L., Zengel J.M.;  
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: Involved in the binding of tRNA to the ribosomes (By  
similarity).  
CC -1- SIMILARITY: BELONGS TO THE S10P FAMILY OF RIBOSOMAL PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AF058451; AAC14294.1; -  
 DR HAMAP; MF\_00508; -; 1.  
 DR InterPro; IPR001848; Ribosomal\_S10.  
 DR PROSITE; PS00361; RIBOSOMAL\_S10; PARTIAL.  
 KW Ribosomal protein.  
 FT NON\_TER 9  
 SQ SEQUENCE 9 AA; 1214 MW; DE3944004416D456 CRC64;

Query Match 31.1%; Score 14; DB 1; Length 9;  
 Best Local Similarity 40.0%; Pred. No. 1.3e+05;  
 Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 6 QERV 10  
 DB 1 1:  
 2 QNORI 6

## RESULT 7

ID MALE KLEPN STANDARD; PRT; 10 AA.  
 AC Q05564;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Maltose-binding periplasmic protein (Maltodextrin-binding protein)  
 DE (MBP) (Fragment).  
 GN MALE.  
 OS Klebsiella pneumoniae.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Klebsiella.  
 OX NCBI\_TaxID=573;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1033-5P14 / KAY2026;  
 RX MEDLINE=93211295; PubMed=8459773;  
 RA Bachelier S., Perrin D., Hofnung M., Gilson E.;  
 RT "Bacterial interspersed mosaic elements (BIMES) are present in the  
 RL genome of Klebsiella."  
 RL Mol. Microbiol. 7:537-544(1993).  
 CC -!- FUNCTION: MALE IS INVOLVED IN THE HIGH-AFFINITY MALTOSE MEMBRANE  
 CC TRANSPORT SYSTEM. INITIAL RECEPTOR FOR THE ACTIVE TRANSPORT OF AND  
 CC CHEMOTAXIS TOWARD MALTOLOGOSACCHARIDES.  
 CC -!- SUBCELLULAR LOCATION: Periplasmic.  
 CC -!- SIMILARITY: BELONGS TO THE BACTERIAL EXTRACELLULAR SOLUTE-BINDING  
 CC PROTEIN FAMILY 1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; X68329; CAA48406.1; -  
 DR InterPro; IPR006061; SBP\_dom1.  
 DR PROSITE; PS01037; SBP\_BACTERIAL\_1; PARTIAL.  
 KW Transport; Sugar transport; Periplasmic.  
 FT NON\_TER 1  
 SQ SEQUENCE 10 AA; 1159 MW; 8FD8DC4415A6DDDA CRC64;

Query Match 31.1%; Score 14; DB 1; Length 10;  
 Best Local Similarity 75.0%; Pred. No. 4.3e+03;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 AQER 8

Db 4 AQSR 7

## RESULT 8

ID UPAL\_HUMAN STANDARD; PRT; 8 AA.  
 AC P30087;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Unknown protein from 2D-page of plasma (Spot 2) (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Plasma;  
 RX MEDLINE=93092937; PubMed=1459097;  
 RA Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,  
 RA Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,  
 RA Hochstrasser D.F.;  
 RT "Plasma protein map: an update by microsequencing."  
 RL Electrophoresis 13:707-714(1992).  
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 CC PROTEIN IS: 4.9, ITS MW IS: 65 kDa.  
 CC SWISS-2DPAGE; P30087; HUMAN.  
 DR NON\_TER 1  
 FT UNSURE 8  
 FT NON\_TER 8  
 SQ SEQUENCE 8 AA; 944 MW; C01772C455BB06DA CRC64;

Query Match 28.9%; Score 13; DB 1; Length 8;  
 Best Local Similarity 60.0%; Pred. No. 1.3e+05;  
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 QERV 10  
 DB 1 1:  
 2 QESNV 6

## RESULT 9

ID Q2OG\_COMTE STANDARD; PRT; 10 AA.  
 AC P80466;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Quinolone 2-oxidoeductase, gamma chain (EC 1.3.99.17) (Fragment).  
 OS Comamonas testosteroni (Pseudomonas testosteroni).  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Comamonadaceae; Comamonas.  
 OX NCBI\_TaxID=285;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=63;  
 RX MEDLINE=96035889; PubMed=7556204;  
 RA Schach S., Tshisuaka B., Fetzner S., Lingens F.;  
 RT "Quinolone 2-oxidoeductase and 2-oxo-1,2-dihydroquinoline 5,6-  
 RT dioxygenase from Comamonas testosteroni 63. The first two enzymes in  
 RT quinoline and 3-methylquinoline degradation."  
 RL Eur. J. Biochem. 232:536-544(1995).  
 CC -!- FUNCTION: CONVERTS (3-METHYL)-QUINOLINE TO (3-METHYL)-2-OXO-  
 CC 1,2-DIHYDROQUINOLINE.  
 CC -!- CATALYTIC ACTIVITY: Quinolone + acceptor + H(2)O = isoquinolin-  
 CC 1(2H)-one + reduced acceptor.  
 CC -!- COFACTOR: FAD, MOLYBDENUM AND IRON-SULFUR.  
 CC -!- PATHWAY: Degradation of quinoline and (3-methyl-)quinoline; first  
 CC step.  
 CC -!- SUBUNIT: HETEROHEXAMER OF TWO ALPHA CHAINS, TWO BETA CHAINS, AND  
 CC TWO GAMMA CHAINS (PROBABLE).  
 KW Oxidoreductase; Flavoprotein; FAD; Molybdenum.

```
FT NON_TER 10 10
SQ SEQUENCE 10 AA; 1153 MW; C848CE64433B1DC6 CRC64;

Query Match
Best Local Similarity 33.3%; Score 13; DB 1; Length 10;
Matches 2; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 MLAAQE 7
   I: I:
   1 MIQAEK 6

Db

RESULT 10
UH11_RAT
ID UH11_RAT STANDARD; PRT; 7 AA.
AC P56576;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Unknown protein from 2D-page of heart tissue (Spot p11) (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE.
RC STRAIN=Wistar; TISSUE=Heart;
RA Li X.-P., Pleissner K.-P., Scheler C., Regitz-Zagrosek V., Salikov J.,
RA Jungblut P.R.;
RL Submitted (SEP-1998) to the SWISS-PROT data bank.
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 8.5, ITS MW IS: 42 kDa.
FT UNSURE 2 2 OR A.
FT NON_TER 7
SQ SEQUENCE 7 AA; 775 MW; 6866DB040DC5A6B0 CRC64;

Query Match
Best Local Similarity 26.7%; Score 12; DB 1; Length 7;
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 AAQER 8
   I: I:
   2 SAREQ 6

Db

RESULT 11
NS3_MYCTU
ID NS3_MYCTU STANDARD; PRT; 8 AA.
AC P81152;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 30 kDa non-secretory protein 3 (Fragment).
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE.
RC STRAIN=H37Rv;
RA Prasad H.K., Annapurna P.S.;
RL Submitted (DEC-1997) to the SWISS-PROT data bank.
CC -1- CAUTION: We are unable to find this protein in the translation of
CC the genome of strain H37Rv.
FT NON_TER 1 1
FT NON_TER 8
SQ SEQUENCE 8 AA; 919 MW; 8D3DC40B19CDC2D2 CRC64;

Query Match
Best Local Similarity 26.7%; Score 12; DB 1; Length 8;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LAAQER 8
```

Db 1 VVAFER 6

#### RESULT 12

FIBB\_CERSI  
ID FIBB\_CERSI STANDARD; PRT; 10 AA.  
AC P14537;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).  
GN FGB.  
OS Ceratotherium simum (White rhinoceros) (Square-lipped rhinoceros).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Ceratotherium.  
OX NCBI\_TaxID=9807;  
RN [1]  
RP SEQUENCE.  
RA O'Neill P.B., Doolittle R.F.;  
RL "Mammalian phylogeny based on fibrinopeptide amino acid sequences.";  
RL Syst. Zool. 22:590-595(1973).  
CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
CC AGGREGATION.  
CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.  
DR InterPro: IPR002181; Fibrinogen\_C.  
DR PROSITE: PS00514; FIBRIN\_AG\_C\_DOMAIN; PARTIAL.  
KW Blood coagulation; Plasma.  
FT PEPTIDE 1 10 FIBRINOPEPTIDE B.  
FT NON\_TER 10  
SQ SEQUENCE 10 AA; 1097 MW; 9402B2B2CDDDD33A CRC64;

Query Match 26.7%; Score 12; DB 1; Length 10;  
Best Local Similarity 50.0%; Pred. No. 1.1e+04;  
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 AAQERR 9  
 I: I:  
 5 AAVDAR 10

#### RESULT 13

GAJU\_HUMAN  
ID GAJU\_HUMAN STANDARD; PRT; 10 AA.  
AC P01358;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Gastric juice peptide.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=75150968; PubMed=5538385;  
RA Heathcote J.G., Washington R.J.;  
RL "Peptides of normal human gastric juice.";  
RL Int. J. Protein Res. 2:117-126(1970).  
DR PIR; A01628; GXHUI.  
DR MIM; I37220; -.  
DR GO; GO:0007586; P:digestion; NAS.  
FT PEPTIDE 1 10 GASTRIC JUICE PEPTIDE 1.  
FT PEPTIDE 2 10 GASTRIC JUICE PEPTIDE 2.  
SQ SEQUENCE 10 AA; 1004 MW; CFEEC6AB02C3387D CRC64;

Query Match 26.7%; Score 12; DB 1; Length 10;

Best Local Similarity 100.0%; Pred. No. 1.1e+04;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAA 5  
   |||  
Db 1 LAA 3

RESULT 14

FARD\_CALVO  
ID FARD\_CALVO STANDARD; PRT; 9 AA.  
AC P41868;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE CalliFMRamide 13.  
OS Calliphora vomitoria (Blue blowfly).  
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;  
OC Calliphoridae; Calliphora.  
OX NCBI\_TaxID=27454;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Thoracic ganglion;  
RX MEDLINE=9219611; PubMed=1549595;  
RA Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,  
RA Rehfeld J.F., Thorpe A.;  
RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2  
RT neuropeptides (designated calliFMRamides) from the blowfly  
RT Calliphora vomitoria.";  
RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).  
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
CC FAMILY.  
DR PIR; D44787.  
KW Neuropeptide; Amidation.  
FT MOD.RES 9 9  
SQ SEQUENCE 9 AA; 1028 MW; 22D10699C87AB6D8 CRC64;

Query Match 24.4%; Score 11; DB 1; Length 9;  
Best Local Similarity 50.0%; Pred. No. 1.3e+05;  
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 AAQE 7  
   |||  
Db 1 AGQD 4

RESULT 15

ULAH\_HUMAN  
ID ULAH\_HUMAN STANDARD; PRT; 9 AA.  
AC P31934;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Unknown protein from 2D-page of liver tissue (Spot 94) (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Liver;  
RX MEDLINE=94147969; PubMed=8313870;  
RA Hughes G.J., Frutiger S., Paquet N., Pasquali C., Sanchez J.-C.,  
RA Tissot J.-D., Balroch A., Appel R.D., Hochstrasser D.F.;  
RT "Human liver protein map: update 1993.";  
RL Electrophoresis 14:1216-1222(1993).  
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
CC PROTEIN IS: 5.5. ITS MW IS: 12 kDa.  
DR SWISS-2DPAGE; P31934; HUMAN.  
FT NON\_TER 1 1  
FT VARIANT 3 3 A -> L.  
FT /FTid=VAR\_000001.

FT NON\_TER 9 9  
SQ SEQUENCE 9 AA; 990 MW; F1CC7B172AB6DDCA CRC64;

Query Match 24.4%; Score 11; DB 1; Length 9;  
Best Local Similarity 66.7%; Pred. No. 1.3e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 AQE 7  
   |||  
Db 3 AGD 5

Search completed: October 7, 2003, 14:28:17  
Job time : 14 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:24:10 ; Search time 58.5 seconds  
(without alignments)  
44.112 Million cell updates/sec

Title: US-09-807-512-25  
Perfect score: 45  
Sequence: 1 AMLAAQERRV 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues  
Total number of hits satisfying chosen parameters: 1349

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : SPTREMBL\_23.\*
- 1: sp\_archaea.\*
  - 2: sp\_bacteria.\*
  - 3: sp\_fungi.\*
  - 4: sp\_human.\*
  - 5: sp\_invertebrate.\*
  - 6: sp\_mammal.\*
  - 7: sp\_mhc.\*
  - 8: sp\_organelle.\*
  - 9: sp\_phage.\*
  - 10: sp\_plant.\*
  - 11: sp\_rodent.\*
  - 12: sp\_virus.\*
  - 13: sp\_vertebrate.\*
  - 14: sp\_unclassified.\*
  - 15: sp\_rvirus.\*
  - 16: sp\_bacteriap.\*
  - 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	17	37.8	9	15 Q85710	Q85710 rous sarcom
2	16	35.6	7	2 Q47477	Q47477 escherichia
3	16	35.6	10	2 P83160	P83160 anabaena sp
4	16	35.6	10	10 Q94119	Q94119 zea mays (m
5	15	33.3	9	4 Q15892	Q15892 homo sapien
6	15	33.3	9	10 Q940K4	Q940K4 arabidopsis
7	15	33.3	9	11 O88889	O88889 mus musculu
8	15	33.3	10	10 O8LR7	O8LR7 chlamydomon
9	15	33.3	10	13 O42355	O42355 brachydanio
10	14	31.1	8	5 O15896	O15896 babesia bov
11	14	31.1	8	9 O8H9K4	O8H9K4 bacterioph
12	14	31.1	8	9 O8H9J9	O8H9J9 bacterioph
13	14	31.1	8	9 O8H9J7	O8H9J7 bacterioph
14	14	31.1	8	9 O8H9J5	O8H9J5 bacterioph
15	14	31.1	8	9 O8H9J3	O8H9J3 bacterioph
16	14	31.1	8	9 O8H9J1	O8H9J1 bacterioph

17	14	31.1	8	9 Q8H9I9	Q8H9I9 bacterioph
18	14	31.1	8	9 Q8H9I8	Q8H9I8 bacterioph
19	14	31.1	8	9 Q8H9I6	Q8H9I6 bacterioph
20	14	31.1	8	9 Q8H9I4	Q8H9I4 bacterioph
21	14	31.1	8	9 Q8H9I3	Q8H9I3 bacterioph
22	14	31.1	8	9 Q8H9I2	Q8H9I2 bacterioph
23	14	31.1	8	9 Q8H9H6	Q8H9H6 bacterioph
24	14	31.1	8	9 Q8H9H5	Q8H9H5 bacterioph
25	14	31.1	8	9 Q8H9H3	Q8H9H3 bacterioph
26	14	31.1	8	9 Q8H9H1	Q8H9H1 bacterioph
27	14	31.1	9	2 Q9K4M6	Q9K4M6 lactococcu
28	14	31.1	9	6 Q9TT77	Q9TT77 bos tauru
29	14	31.1	9	6 Q9TRW2	Q9TRW2 oryctolagu
30	14	31.1	10	1 Q9UWM5	Q9UWM5 sulfolobu
31	14	31.1	10	4 Q9H3R9	Q9H3R9 homo sapien
32	14	31.1	10	10 P82132	P82132 spinacia ol
33	14	31.1	10	10 P82133	P82133 spinacia ol
34	14	31.1	10	11 Q9QV21	Q9QV21 rattus sp.
35	14	31.1	10	12 P90373	P90373 pseudorabie
36	14	31.1	10	13 O73588	O73588 gallus gall
37	13	28.9	7	8 P92214	P92214 amblyopyrum
38	13	28.9	7	8 P92393	P92393 hordeum vul
39	13	28.9	7	8 P92403	P92403 lophopyrum
40	13	28.9	7	8 P92427	P92427 peridictyon
41	13	28.9	7	8 P92430	P92430 aegilops ta
42	13	28.9	7	8 P92221	P92221 bromus iner
43	13	28.9	7	8 P92425	P92425 pseudoroegn
44	13	28.9	7	8 P92381	P92381 hordeum bra
45	13	28.9	7	8 P92387	P92387 henrardia p

ALIGNMENTS

RESULT 1

Q85710 ID Q85710 PRELIMINARY; PRT; 9 AA.  
AC Q85710;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
DE Pol protein (Fragment).  
OS Rous sarcoma virus.  
OC Viruses; Retroid viruses; Retroviridae; Alpharetrovirus.  
OX NCBI\_TaxID=11886;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=84115080; PubMed=6319754;  
RA Lerner T.L., Hanafusa H.;  
RT "DNA sequence of the Bryan high-titer strain of Rous sarcoma virus:  
RT Extent of env deletion and possible genealogical relationship with  
RT other viral strains.";  
RL J. Virol. 49:549-556(1984).  
DR EMBL; K03365; AAA42557.1; -.  
FT NON\_TER 1  
SQ SEQUENCE 9 AA; 949 MW; 94AA144DDDD731AA CRC64;

Query Match 37.8%; Score 17; DB 15; Length 9;  
Best Local Similarity 80.0%; Pred. No. 8.3e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 LAAQ 7  
Db 4 LAANE 8

RESULT 2

Q47477 ID Q47477 PRELIMINARY; PRT; 7 AA.  
AC Q47477;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)

```

DE Tpi protein (Fragment).
GN TPI.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE OF 7-7 FROM N.A.
RX MEDLINE=85203917; PubMed=3158524;
RA Hellings H.W., Evans P.R.;
RT "Nucleotide sequence and high-level expression of the major
RT Escherichia coli phosphofructokinase.";
RL Eur. J. Biochem. 149:363-373(1985).
RN [2]
RP SEQUENCE FROM N.A.
RA Evans P.;
RL Submitted (OCT-1986) to the EMBL/GenBank/DBJ databases.
DR EMBL; X02519; CAA26359.1; -.
FT NON_TER 1
SQ SEQUENCE 7 AA; 773 MW; 7416D3DDDB1DB0 CRC64;

Query Match 35.6%; Score 16; DB 2; Length 7;
Best Local Similarity 60.0%; Pred. No. 8.3e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 AAQR 8
Db 3 AAQR 7

RESULT 3
P83160 PRELIMINARY; PRT; 10 AA.
ID P83160;
AC P83160;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Ribulose biphosphate carboxylase small chain (EC 4.1.1.39) (RuBisCO
DE small subunit) (Fragment).
OS Anabaena sp. (strain L31).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.
OX NCBI_TaxID=29412;
RN [1]
RP SEQUENCE.
RA Apte S.K., Uhlemann E., Schmid R., Altendorf K.;
RL Submitted (OCT-2001) to the SWISS-PROT data bank.
CC -1- FUNCTION: RUBISCO CATALYZES TWO REACTIONS: THE CARBOXYLATION OF D-
CC RIBULOSE 1,5-BISPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC
CC CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF
CC THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH
CC REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME
CC ACTIVE SITE.
CC -1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + CO(2) = 2 3-
CC PHOSPHO-D-GLYCERATE.
CC -1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + O(2) = 3-
CC PHOSPHO-D-GLYCERATE + 2-PHOSPHOGLYCOLATE.
CC -1- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE RUBISCO SMALL CHAIN FAMILY.
KW Photosynthesis; Carbon dioxide fixation; Photorepiration; Lyase;
KW Oxidoreductase; Monooxygenase; Calvin cycle.
FT NON_TER 10
SQ SEQUENCE 10 AA; 1322 MW; 8B4E2D0B13276731 CRC64;

Query Match 35.6%; Score 16; DB 2; Length 10;
Best Local Similarity 75.0%; Pred. No. 1.1e+04;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 QERR 9
Db 6 KERR 9

```

```

RESULT 4
Q941I9 PRELIMINARY; PRT; 10 AA.
ID Q941I9;
AC Q941I9;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Anthocyanin regulator R-sc protein (Fragment).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RA Prociassi A., Piazza P., Tonelli C.;
RT "A maize rl gene is regulated post-transcriptionally by differential
RT splicing of its leader.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY029766; AAK61356.1; -.
FT NON_TER 10
SQ SEQUENCE 10 AA; 1074 MW; 9CC8E4DDDD736C5 CRC64;

Query Match 35.6%; Score 16; DB 10; Length 10;
Best Local Similarity 50.0%; Pred. No. 1.1e+04;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AMLAAQER 8
Db 3 SQAAR 10

RESULT 5
Q15892 PRELIMINARY; PRT; 9 AA.
ID Q15892;
AC Q15892;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE (Clone XP3B4A) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Placenta;
RA Lee C.-C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,
RA Coolbaugh M.I., Chinault C.A., Baldini A., Lindsay E.A., Zhao Z.-Y.,
RA Caskey C.T.H.;
RT "Isolation of chromosome-specific genes by reciprocal probing of
RT arrayed cDNAs and cosmid libraries.";
RL Hum. Mol. Genet. 0:0-0(1995).
DR EMBL; L32071; AAA73882.1; -.
FT NON_TER 1
FT NON_TER 9
SQ SEQUENCE 9 AA; 971 MW; 49B22732CDC40B17 CRC64;

Query Match 33.3%; Score 15; DB 4; Length 9;
Best Local Similarity 66.7%; Pred. No. 8.3e+05;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 AQERV 10
Db 1 ALERAV 6

RESULT 6
Q940K4 PRELIMINARY; PRT; 9 AA.
ID Q940K4;
AC Q940K4;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)

```

DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
DE Hypothetical 1.4 kDa protein.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,  
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,  
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,  
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,  
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,  
RA Ecker J., Theologis A., Davis R.W.;  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,  
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,  
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,  
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,  
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,  
RA Ecker J., Theologis A., Davis R.W.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY054502; AAK96693.1; -;  
DR EMBL; AY081672; AAM10234.1; -;  
KW Hypothetical protein.  
SQ SEQUENCE 9 AA; 1362 MW; 712D1416D9D41414 CRC64;  
  
Query Match 33.3%; Score 15; DB 10; Length 9;  
Best Local Similarity 75.0%; Pred. No. 8.3e+05;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 6 QERR 9  
DB 6 QRRR 9  
  
RESULT 7  
O88889 PRELIMINARY; PRT; 9 AA.  
AC O88889;  
DT 01-NOV-1998 (TReMBLrel. 08, Created)  
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE Ubiquitin-conjugating enzyme UBCM4 (Fragment).  
GN UBCM4.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129/SV;  
RX MEDLINE=97057256; PubMed=8901595;  
RT "Protein integration into a gene encoding a ubiquitin-conjugating  
RT enzyme results in a placental defect and embryonic lethality";  
RL Proc. Natl. Acad. Sci. U.S.A. 93:12412-12417(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129/SV;  
RX MEDLINE=99132641;  
RA Miller U., Grams A., Martinez-Noel G., Copeland N.G., Gilbert D.J.,  
RA Jenkins N.A., Harbers K.;  
RT "Structure of the gene encoding the ubiquitin-conjugating enzyme  
RT UbcM4, characterization of its promoter, and chromosomal location";  
RL Gene 224:109-116(1998).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129/SV;  
RA Mueller U., Grams A., Martinez-Noel G., Harbers K.;

RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF071557; AAD10128.1; -;  
FT NON\_TER 9  
SQ SEQUENCE 9 AA; 1063 MW; C90F97341415BDDC CRC64;  
  
Query Match 33.3%; Score 15; DB 11; Length 9;  
Best Local Similarity 50.0%; Pred. No. 8.3e+05;  
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
QY 3 LAAQER 8  
DB 1 MAASRR 6  
  
RESULT 8  
O8LRT7 PRELIMINARY; PRT; 10 AA.  
AC O8LRT7;  
DT 01-OCT-2002 (TReMBLrel. 22, Created)  
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)  
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)  
DE Zyslb (Fragment)  
OS Chlamydomonas reinhardtii.  
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
OC Chlamydomonadaceae; Chlamydomonas.  
OX NCBI\_TaxID=3055;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CC-2290;  
RA Kathir P., LaVoie M., Lefebvre P.A., Silflow C.D.;  
RT "Molecular Map of the Chlamydomonas Nuclear Genome";  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF503638; AAM27452.1; -;  
FT NON\_TER 1  
SQ SEQUENCE 10 AA; 1071 MW; C7D8D8C6C6DC6C6 CRC64;  
  
Query Match 33.3%; Score 15; DB 10; Length 10;  
Best Local Similarity 66.7%; Pred. No. 1.8e+04;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 AMLAAQ 6  
DB 5 AQAQAQ 10  
  
RESULT 9  
O42355 PRELIMINARY; PRT; 10 AA.  
AC O42355;  
DT 01-JAN-1998 (TReMBLrel. 05, Created)  
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)  
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)  
DE GATA-2 (Fragment).  
GN GATA2.  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=9732361; PubMed=9177206;  
RA Meng A., Tang H., Ong B.A., Farrell M.J., Lin S.;  
RT "Promoter analysis in living embryos identifies a cis-acting motif  
RL required for neuronal expression of GATA-2";  
RL Proc. Natl. Acad. Sci. U.S.A. 94:6267-6272(1997).  
DR EMBL; AF001220; AAB61711.1; -;  
DR ZFIN; ZDB-GENE-980526-260; gata2.  
FT NON\_TER 10  
SQ SEQUENCE 10 AA; 1192 MW; C82A2CA6DAADDC2 CRC64;  
  
Query Match 33.3%; Score 15; DB 13; Length 10;  
Best Local Similarity 42.9%; Pred. No. 1.8e+04;

Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 LAQAQRR 9  
Db 3 VAAQSR 9

## RESULT 10

O15896  
ID O15896 PRELIMINARY; PRT; 8 AA.  
AC O15896;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE 12D3 antigen (Fragment).  
GN 12D3.  
OS Babesia bovis.  
OC Eukaryota; Alveolata; Apicomplexa; Piroplasmida; Babesiidae; Babesia.  
OX NCBI\_TaxID=5865;  
RN [1]  
RC STRAIN=Samford attenuated;  
RA Silins G.U., Blakeley R.L., Riddles P.W.;  
RT "Characterization of a gene encoding an ATP-binding protein from Babesia bovis."  
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RC STRAIN=Samford attenuated;  
RA Silins G.U., Blakeley R.L., Riddles P.W.;  
RT "Characterization of the transcriptional control region of the 12D3 antigen gene from the sporozoan Babesia bovis."  
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RC STRAIN=Samford attenuated;  
RA Silins G.U., Blakeley R.L., Riddles P.W.;  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U44917; AAB66362.1; -.  
DR EMBL; AY170919; AAN64587.1; -.  
FT NON\_TER 8  
SQ SEQUENCE 8 AA; 984 MW; F0D2C9D411ADD726 CRC64;

Query Match 31.1%; Score 14; DB 5; Length 8;  
Best Local Similarity 60.0%; Pred. No. 8.3e+05;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 MLAAQ 6  
Db 1 MLATR 5

## RESULT 11

O8H9K4  
ID O8H9K4 PRELIMINARY; PRT; 8 AA.  
AC O8H9K4;  
DT 01-MAR-2003 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Gpvs.1 protein (Fragment).  
GN VS.1.  
OS Bacteriophage K3.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.  
OX NCBI\_TaxID=10674;  
RN [1]  
RC SEQUENCE FROM N.A.  
RA Plesiniene L., Nivinskis R.;  
RT "Analysis of sequence-specific endoribonuclease RegB gene in T4-related bacteriophages."  
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ488515; CAD32707.1; -.  
FT NON\_TER 8  
SQ SEQUENCE 8 AA; 859 MW; F17DD7272DD32406 CRC64;

Query Match 31.1%; Score 14; DB 9; Length 8;  
Best Local Similarity 75.0%; Pred. No. 8.3e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AMLA 4  
Db 4 ALLA 7

## RESULT 12

O8H9J9  
ID O8H9J9 PRELIMINARY; PRT; 8 AA.  
AC O8H9J9;  
DT 01-MAR-2003 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Gpvs.1 protein (Fragment).  
GN VS.1.  
OS Bacteriophage M1.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.  
OX NCBI\_TaxID=10676;  
RN [1]  
RC SEQUENCE FROM N.A.  
RA Plesiniene L., Nivinskis R.;  
RT "Analysis of sequence-specific endoribonuclease RegB gene in T4-related bacteriophages."  
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ488514; CAD32705.1; -.  
FT NON\_TER 8  
SQ SEQUENCE 8 AA; 859 MW; F17DD7272DD32406 CRC64;

Query Match 31.1%; Score 14; DB 9; Length 8;  
Best Local Similarity 75.0%; Pred. No. 8.3e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AMLA 4  
Db 4 ALLA 7

## RESULT 13

O8H9J7  
ID O8H9J7 PRELIMINARY; PRT; 8 AA.  
AC O8H9J7;  
DT 01-MAR-2003 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Gpvs.1 protein (Fragment).  
GN VS.1.  
OS Bacteriophage Ox2.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;  
OX NCBI\_TaxID=10691;  
RN [1]  
RC SEQUENCE FROM N.A.  
RA Plesiniene L., Nivinskis R.;  
RT "Analysis of sequence-specific endoribonuclease RegB gene in T4-related bacteriophages."  
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ488516; CAD32709.1; -.  
FT NON\_TER 8  
SQ SEQUENCE 8 AA; 859 MW; F17DD7272DD32406 CRC64;

Query Match 31.1%; Score 14; DB 9; Length 8;  
Best Local Similarity 75.0%; Pred. No. 8.3e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AMLA 4  
Db 4 ALLA 7

## RESULT 14

```
Q8H9J5
ID Q8H9J5 PRELIMINARY; PRT; 8 AA.
AC Q8H9J5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Gpvs.1 protein (Fragment).
GN VS.1.
OS Bacteriophage RB2.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC T4-like viruses.
OX NCBI_TaxID=164992;
RN [1]
RP SEQUENCE FROM N.A.
RA Piesiniene L., Nivinskas R.;
RT "Analysis of sequence-specific endoribonuclease RegB gene in T4-
RT related bacteriophages.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ488519; CAD32715.1; -.
FT NON_TER 8
SQ SEQUENCE 8 AA; 859 MW; F17DD7272DD32406 CRC64;

Query Match 31.1%; Score 14; DB 9; Length 8;
Best Local Similarity 75.0%; Pred. No. 8.3e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AMLA 4
DB 4 ALLA 7

RESULT 15
Q8H9J3
ID Q8H9J3 PRELIMINARY; PRT; 8 AA.
AC Q8H9J3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Gpvs.1 protein (Fragment).
GN VS.1.
OS Bacteriophage RB3.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC T4-like viruses.
OX NCBI_TaxID=10695;
RN [1]
RP SEQUENCE FROM N.A.
RA Piesiniene L., Nivinskas R.;
RT "Analysis of sequence-specific endoribonuclease RegB gene in T4-
RT related bacteriophages.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ488520; CAD32717.1; -.
FT NON_TER 8
SQ SEQUENCE 8 AA; 859 MW; F17DD7272DD32406 CRC64;

Query Match 31.1%; Score 14; DB 9; Length 8;
Best Local Similarity 75.0%; Pred. No. 8.3e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AMLA 4
DB 4 ALLA 7

Search completed: October 7, 2003, 14:30:23
Job time : 59.5 secs
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:25:50 ; Search time 19.5 Seconds  
(without alignments)  
21.698 Million cell updates/sec

Title: US-09-807-512-25

Perfect score: 45

Sequence: 1 AMLAQRRV 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 90058

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*

2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*

3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*

4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*

5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*

6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	55.6	9	3	US-08-144-779C-27
2	25	55.6	9	4	US-09-368-449B-27
3	22	48.9	7	3	US-08-144-779C-24
4	22	48.9	7	4	US-09-368-449B-24
5	22	48.9	8	3	US-08-144-779C-25
6	22	48.9	8	4	US-09-368-449B-25
7	22	48.9	9	3	US-08-144-779C-26
8	22	48.9	9	4	US-09-368-449B-26
9	22	48.9	9	4	US-09-447-399-7
10	21	46.7	6	4	US-08-055-006-53
11	21	46.7	7	4	US-09-627-851B-39
12	21	46.7	9	3	US-09-041-886-45
13	21	46.7	9	3	US-09-518-046-78
14	21	46.7	9	3	US-09-518-046-95
15	21	46.7	9	3	US-09-518-046-146
16	21	46.7	9	4	US-09-644-600-59
17	20	44.4	6	3	US-08-463-160B-65
18	20	44.4	7	1	US-08-552-142A-5
19	20	44.4	7	1	US-08-910-973-5
20	20	44.4	7	4	US-09-499-227-5
21	20	44.4	7	5	PCT-US95-05741-5
22	20	44.4	8	3	US-08-144-779C-31
23	20	44.4	8	4	US-09-368-449B-31
24	20	44.4	9	3	US-08-159-339A-224
25	20	44.4	9	4	US-09-434-476A-28
26	20	44.4	9	4	US-09-434-476A-29
27	20	44.4	10	1	US-08-336-343A-40

28 19 42.2 5 4 US-09-638-202A-63 Sequence 63, Appl  
29 19 42.2 6 2 US-08-318-837-22 Sequence 22, Appl  
30 19 42.2 10 2 US-08-292-968-45 Sequence 45, Appl  
31 19 42.2 10 2 US-08-467-974-45 Sequence 45, Appl  
32 19 42.2 10 2 US-08-467-974-45 Sequence 45, Appl  
33 19 42.2 10 3 US-08-467-976-45 Sequence 45, Appl  
34 19 42.2 10 3 US-08-467-976-45 Sequence 45, Appl  
35 18 40.0 6 2 US-09-082-514-45 Sequence 17, Appl  
36 18 40.0 7 1 US-07-820-154A-25 Sequence 25, Appl  
37 18 40.0 7 2 US-08-097-554A-25 Sequence 25, Appl  
38 18 40.0 7 3 US-08-480-640A-25 Sequence 25, Appl  
39 18 40.0 7 3 US-09-413-452-48 Sequence 25, Appl  
40 18 40.0 7 3 US-08-295-802-25 Sequence 25, Appl  
41 18 40.0 7 3 US-08-488-237A-25 Sequence 25, Appl  
42 18 40.0 7 3 US-09-413-068-48 Sequence 48, Appl  
43 18 40.0 7 4 US-08-375-992A-25 Sequence 25, Appl  
44 18 40.0 7 4 US-08-472-679H-25 Sequence 25, Appl  
45 18 40.0 7 5 PCT-US93-00324-25 Sequence 25, Appl

#### ALIGNMENTS

RESULT 1  
US-08-144-779C-27  
; Sequence 27, Application US/08144779C  
; Patent No. 6066622  
; GENERAL INFORMATION:  
; APPLICANT: Green, Lawrence R.  
; APPLICANT: Sinackevich, Nicolay V.  
; APPLICANT: Ivanov, Vadim T.  
; APPLICANT: Mikhailova, Inessa I.  
; APPLICANT: Vaskovsky, Boris V.  
; APPLICANT: Mikhailov, Alexander N.  
; APPLICANT: Khavinson, Vladimir K.  
; APPLICANT: Morozov, Vyacheslav G.  
; TITLE OF INVENTION: Immunomodulating Peptides and Methods  
; TITLE OF INVENTION: of Use  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/144,779C  
; FILING DATE: 28-OCT-1993  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/783,517  
; FILING DATE: 28-OCT-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/816,205  
; FILING DATE: 02-JAN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/967,633  
; FILING DATE: 28-OCT-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Storella, John R.  
; REGISTRATION NUMBER: 32,944  
; REFERENCE/DOCKET NUMBER: 15542-5-2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:

;  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-144-779C-27

Query Match 55.6%; Score 25; DB 3; Length 9;  
Best Local Similarity 44.4%; Pred. No. 2.5e+05;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 MLAAQRRV 10  
|||::|  
Db 1 MLTAZZKAV 9

## RESULT 2

US-09-368-449B-27  
; Sequence 27, Application US/09368449B  
; Patent No. 6346514

## GENERAL INFORMATION:

APPLICANT: Green, Lawrence R.  
Sinackevich, Nicolay V.  
Ivanov, Vadim T.  
Mikhalyova, Inessa I.  
Vaskovsky, Boris V.  
Mikhaltsov, Alexander N.  
Khavinson, Vladimir K.  
Morozov, Vyacheslav G.

TITLE OF INVENTION: Immunomodulating Peptides and Methods  
Of Use

NUMBER OF SEQUENCES: 42

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/368,449B

FILING DATE: 04-Aug-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/144,779

FILING DATE: <Unknown>

APPLICATION NUMBER: US 07/816,205

FILING DATE: 02-JAN-1992

APPLICATION NUMBER: US 07/967,633

FILING DATE: 28-OCT-1992

ATTORNEY/AGENT INFORMATION:

NAME: Storella, John R.

REGISTRATION NUMBER: 32,944

REFERENCE/DOCKET NUMBER: 15542-5-2

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:

LENGTH: 9 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 27:

US-09-368-449B-27

Query Match

55.6%; Score 25; DB 4; Length 9;

Best Local Similarity 44.4%; Pred. No. 2.5e+05;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 MLAAQRRV 10  
|||::|  
Db 1 MLTAZZKAV 9

## RESULT 3

US-08-144-779C-24

; Sequence 24, Application US/08144779C

; Patent No. 6066622

; GENERAL INFORMATION:

APPLICANT: Green, Lawrence R.  
Sinackevich, Nicolay V.  
Ivanov, Vadim T.  
Mikhalyova, Inessa I.  
Vaskovsky, Boris V.  
Mikhaltsov, Alexander N.  
APPLICANT: Khavinson, Vladimir K.  
Morozov, Vyacheslav G.

TITLE OF INVENTION: Immunomodulating Peptides and Methods  
Of Use

NUMBER OF SEQUENCES: 42

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/144,779C

FILING DATE: 28-OCT-1993

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/783,517

FILING DATE: 28-OCT-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/816,205

FILING DATE: 02-JAN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/967,633

FILING DATE: 28-OCT-1992

ATTORNEY/AGENT INFORMATION:

NAME: Storella, John R.

REGISTRATION NUMBER: 32,944

REFERENCE/DOCKET NUMBER: 15542-5-2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 24:

SEQUENCE CHARACTERISTICS:

LENGTH: 7 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-144-779C-24

Query Match

48.9%; Score 22; DB 3; Length 7;

Best Local Similarity 42.9%; Pred. No. 2.5e+05;

Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 MLAAQER 8

|||::|

Db 1 MLTAZZK 7



```
RESULT 4
US-09-368-449B-24
; Sequence 24, Application US/09368449B
; Patent No. 6346514
; GENERAL INFORMATION:
; APPLICANT: Green, Lawrence R.
; Sinackevich, Nicolay V.
; Ivanov, Vadim T.
; Mikhailova, Inessa I.
; Vaskovsky, Boris V.
; Mikhailsov, Alexander N.
; Khavinson, Vladimir K.
; Morozov, Vyacheslav G.
; TITLE OF INVENTION: Immunomodulating Peptides and Methods
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/368,449B
; FILING DATE: 04-Aug-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/144,779
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/816,205
; FILING DATE: 02-JAN-1992
; APPLICATION NUMBER: US 07/967,633
; FILING DATE: 28-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Storella, John R.
; REGISTRATION NUMBER: 32,944
; REFERENCE/DOCKET NUMBER: 15542-5-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-09-368-449B-24

Query Match 48.9%; Score 22; DB 4; Length 7;
Best Local Similarity 42.9%; Pred. No. 2.5e+05;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 MLAQER 8
Db 1 MLTAZK 7

RESULT 5
US-08-144-779C-25
; Sequence 25, Application US/08144779C
; Patent No. 6066622
; GENERAL INFORMATION:
; APPLICANT: Green, Lawrence R.
; Sinackevich, Nicolay V.
```

```
; APPLICANT: Ivanov, Vadim T.
; APPLICANT: Mikhailova, Inessa I.
; APPLICANT: Vaskovsky, Boris V.
; APPLICANT: Mikhailsov, Alexander N.
; APPLICANT: Khavinson, Vladimir K.
; APPLICANT: Morozov, Vyacheslav G.
; TITLE OF INVENTION: Immunomodulating Peptides and Methods
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/144,779C
; FILING DATE: 28-OCT-1993
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/783,517
; FILING DATE: 28-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/816,205
; FILING DATE: 02-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/967,633
; FILING DATE: 28-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Storella, John R.
; REGISTRATION NUMBER: 32,944
; REFERENCE/DOCKET NUMBER: 15542-5-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-144-779C-25

Query Match 48.9%; Score 22; DB 3; Length 8;
Best Local Similarity 42.9%; Pred. No. 2.5e+05;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 MLAQER 8
Db 1 MLTAZK 7

RESULT 6
US-09-368-449B-25
; Sequence 25, Application US/09368449B
; Patent No. 6346514
; GENERAL INFORMATION:
; APPLICANT: Green, Lawrence R.
; Sinackevich, Nicolay V.
; Ivanov, Vadim T.
; Mikhailova, Inessa I.
; Vaskovsky, Boris V.
; Mikhailsov, Alexander N.
; Khavinson, Vladimir K.
; Morozov, Vyacheslav G.
; TITLE OF INVENTION: Immunomodulating Peptides and Methods
```

```
;
;
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/368,449B
; FILING DATE: 04-Aug-1999
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/144,779
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/816,205
; FILING DATE: 02-JAN-1992
; APPLICATION NUMBER: US 07/967,633
; FILING DATE: 28-OCT-1992
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Storella, John R.
; REGISTRATION NUMBER: 32,944
; REFERENCE/DOCKET NUMBER: 15542-5-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
;
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-368-449B-25

Query Match 48.98; Score 22; DB 4; Length 8;
Best Local Similarity 42.9%; Pred. No. 2.5e+05;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 MAAQER 8
Db 1 MLTAZZK 7

RESULT 7
US-08-144-779C-26
; Sequence 26, Application US/08144779C
; Patent No. 6066622
; GENERAL INFORMATION:
; APPLICANT: Green, Lawrence R.
; APPLICANT: Sinackevich, Nicolay V.
; APPLICANT: Ivanov, Vadim T.
; APPLICANT: Mikhailova, Inessa I.
; APPLICANT: Vaskovsky, Boris V.
; APPLICANT: Mikhailov, Alexander N.
; APPLICANT: Morozov, Vladimir K.
; APPLICANT: Morozov, Vyacheslav G.
; TITLE OF INVENTION: Immunomodulating Peptides and Methods
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
```

```
;
;
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/144,779C
; FILING DATE: 28-OCT-1993
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/783,517
; FILING DATE: 28-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/816,205
; FILING DATE: 02-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/967,633
; FILING DATE: 28-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Storella, John R.
; REGISTRATION NUMBER: 32,944
; REFERENCE/DOCKET NUMBER: 15542-5-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
;
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-144-779C-26

Query Match 48.9%; Score 22; DB 3; Length 9;
Best Local Similarity 42.9%; Pred. No. 2.5e+05;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 MAAQER 8
Db 1 MLTAZZK 7

RESULT 8
US-09-368-449B-26
; Sequence 26, Application US/09368449B
; Patent No. 6346514
; GENERAL INFORMATION:
; APPLICANT: Green, Lawrence R.
; APPLICANT: Sinackevich, Nicolay V.
; APPLICANT: Ivanov, Vadim T.
; APPLICANT: Mikhailova, Inessa I.
; APPLICANT: Vaskovsky, Boris V.
; APPLICANT: Mikhailov, Alexander N.
; APPLICANT: Morozov, Vladimir K.
; APPLICANT: Morozov, Vyacheslav G.
; TITLE OF INVENTION: Immunomodulating Peptides and Methods
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

; APPLICATION NUMBER: US/09/368,449B  
; FILING DATE: 04-Aug-1999  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/144,779  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: US 07/816,205  
; FILING DATE: 02-JAN-1992  
; APPLICATION NUMBER: US 07/967,633  
; FILING DATE: 28-OCT-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Storella, John R.  
; REGISTRATION NUMBER: 32,944  
; REFERENCE/DOCKET NUMBER: 15542-5-2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 26:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: <Unknown>  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 26:  
US-09-368-449B-26

Query Match 48.9%; Score 22; DB 4; Length 9;  
Best Local Similarity 42.9%; Pred. No. 2.5e+05;  
Matches 3; Conservative 3; Mismatches 1; Indels 0;

Qy 2 MAAQER 8  
|||:  
Db 1 MLTAZK 7

RESULT 9  
US-09-447-399-7  
; Sequence 7, Application US/09447399  
; Patent No. 6525187  
; GENERAL INFORMATION:  
; APPLICANT: El Shami, Said A.  
; APPLICANT: Menon, Surendra N.  
; APPLICANT: French, Cynthia K.  
; TITLE OF INVENTION: Polynucleotide Encoding Autoantigens Associated With  
; FILE OF INVENTION: Endometriosis  
; FILE REFERENCE: 107-208US-CIP  
; CURRENT APPLICATION NUMBER: US/09/447,399  
; CURRENT FILING DATE: 1999-11-23  
; EARLIER APPLICATION NUMBER: 60/094,930  
; EARLIER FILING DATE: 1998-07-31  
; EARLIER APPLICATION NUMBER: 09/359,084  
; EARLIER FILING DATE: 1999-07-22  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-447-399-7

Query Match 48.9%; Score 22; DB 4; Length 9;  
Best Local Similarity 66.7%; Pred. No. 2.5e+05;  
Matches 4; Conservative 2; Mismatches 0; Indels 0;

Qy 4 AAQERR 9  
:|:|:  
Db 4 SAEERR 9

RESULT 10  
US-08-055-006-53  
; Sequence 53, Application US/08055006

; Patent No. 6464978  
; GENERAL INFORMATION:  
; APPLICANT: Brostoff, Steven W.  
; Wilson, Darcy B.  
; Smith, Lawrence R.  
; Gold, Daniel P.  
; Carlo, Dennis J.  
; TITLE OF INVENTION: Vaccination and Methods Against Multiple  
; Sclerosis Resulting From Pathogenic Responses By Specific  
; Cell Populations  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CAMPBELL & FLORES LLP  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: United States  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/055,006  
; FILING DATE: 09-Feb-1993  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-1M 9611  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619)535-9001  
; TELEFAX: (619)535-8949  
; INFORMATION FOR SEQ ID NO: 53:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 53:  
US-08-055-006-53

Query Match 46.7%; Score 21; DB 4; Length 6;  
Best Local Similarity 80.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0;

Qy 6 QRRV 10  
|:|:  
Db 1 QDRRV 5

RESULT 11  
US-09-627-851B-39  
; Sequence 39, Application US/09627851B  
; Patent No. 6545131  
; GENERAL INFORMATION:  
; APPLICANT: ISSACS, JOHN T.  
; APPLICANT: DENMEADE, SAMUEL R.  
; APPLICANT: CHRISTENSEN, S. BROGGER  
; APPLICANT: LILJA, HANS  
; TITLE OF INVENTION: TISSUE SPECIFIC PRODRUG  
; FILE REFERENCE: 57111 (71699)  
; CURRENT APPLICATION NUMBER: US/09/627,851B  
; CURRENT FILING DATE: 2000-07-28  
; PRIOR APPLICATION NUMBER: 09/588,822  
; PRIOR FILING DATE: 2000-06-07  
; PRIOR APPLICATION NUMBER: 09/081,707  
; PRIOR FILING DATE: 1998-05-19  
; PRIOR APPLICATION NUMBER: 60/047,070  
; PRIOR FILING DATE: 1997-05-19  
; PRIOR APPLICATION NUMBER: 60/080,046  
; PRIOR FILING DATE: 1998-03-30

; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 39  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-627-851B-39

Query Match 46.7%; Score 21; DB 4; Length 7;  
Best Local Similarity 66.7%; Pred. No. 2.5e+05;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 AQERRV 10  
Db 2 AQRRRL 7

## RESULT 12

US-09-041-886-45  
; Sequence 45, Application US/09041886  
; Patent No. 6235872  
; GENERAL INFORMATION:  
; APPLICANT: Bredesen, Dale E.  
; TITLE OF INVENTION: Protoproptic Peptides, Dependence  
; TITLE OF INVENTION: Polypeptides and Methods of Use  
; NUMBER OF SEQUENCES: 72  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell & Flores LLP  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: United States  
; ZIP: 92122

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/041,886  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LJ 2626  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 45:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-041-886-45

Query Match 46.7%; Score 21; DB 3; Length 9;  
Best Local Similarity 62.5%; Pred. No. 2.5e+05;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AMLAAQR 8  
Db 1 ALLAAALRR 8

## RESULT 13

US-09-518-046-78  
; Sequence 78, Application US/09518046  
; Patent No. 6294663  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.

; APPLICANT: Underwood, Lowell J.  
; TITLE OF INVENTION: Transmembrane Serine Protease Overexpressed  
; TITLE OF INVENTION: in Ovarian Carcinoma and Uses Thereof  
; FILE REFERENCE: D6192CIP  
; CURRENT APPLICATION NUMBER: US/09/518,046  
; CURRENT FILING DATE: 2000-03-02  
; EARLIER APPLICATION NUMBER: 09/261,416  
; EARLIER FILING DATE: 1999-03-03  
; NUMBER OF SEQ ID NOS: 153  
; SEQ ID NO 78  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; LOCATION: 406...414  
; OTHER INFORMATION: TADG-12 peptide  
US-09-518-046-78

Query Match 46.7%; Score 21; DB 3; Length 9;  
Best Local Similarity 80.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 QERRV 10  
Db 3 QERRL 7

## RESULT 14

US-09-518-046-95  
; Sequence 95, Application US/09518046  
; Patent No. 6294663  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Underwood, Lowell J.  
; TITLE OF INVENTION: Transmembrane Serine Protease Overexpressed  
; TITLE OF INVENTION: in Ovarian Carcinoma and Uses Thereof  
; FILE REFERENCE: D6192CIP  
; CURRENT APPLICATION NUMBER: US/09/518,046  
; CURRENT FILING DATE: 2000-03-02  
; EARLIER APPLICATION NUMBER: 09/261,416  
; EARLIER FILING DATE: 1999-03-03  
; NUMBER OF SEQ ID NOS: 153  
; SEQ ID NO 95  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; LOCATION: 407...415  
; OTHER INFORMATION: TADG-12 peptide  
US-09-518-046-95

Query Match 46.7%; Score 21; DB 3; Length 9;  
Best Local Similarity 80.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 QERRV 10  
Db 2 QERRL 6

## RESULT 15

US-09-518-046-146  
; Sequence 146, Application US/09518046  
; Patent No. 6294663  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Underwood, Lowell J.  
; TITLE OF INVENTION: Transmembrane Serine Protease Overexpressed  
; TITLE OF INVENTION: in Ovarian Carcinoma and Uses Thereof  
; FILE REFERENCE: D6192CIP  
; CURRENT APPLICATION NUMBER: US/09/518,046  
; CURRENT FILING DATE: 2000-03-02  
; EARLIER APPLICATION NUMBER: 09/261,416

; EARLIER FILING DATE: 1999-03-03  
; NUMBER OF SEQ ID NOS: 153  
; SEQ ID NO 146  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; LOCATION: 408...416  
; OTHER INFORMATION: TADG-12 peptide  
US-09-518-046-146

Query Match 46.7%; Score 21; DB 3; Length 9;  
Best Local Similarity 80.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 6 QRRV 10  
Db 1 QRRRL 5

Search completed: October 7, 2003, 14:32:04  
Job time : 20.5 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:30:31 ; Search time 41 Seconds  
(without alignments)  
38.589 Million cell updates/sec

Title: US-09-807-512-25  
Perfect score: 45  
Sequence: 1 AMLAAQRRV 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues

Total number of hits satisfying chosen parameters: 77217

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23	51.1	9	11	US-09-799-250-40
2	23	51.1	9	11	US-09-799-250-56
3	23	51.1	9	11	US-09-799-250-121
4	23	51.1	9	11	US-09-799-250-244
5	23	51.1	9	11	US-09-799-250-348
6	23	51.1	9	11	US-09-799-250-560
7	23	51.1	10	11	US-09-799-250-93
8	23	51.1	10	11	US-09-799-250-178
9	23	51.1	10	11	US-09-799-250-185
10	23	51.1	10	11	US-09-799-250-274
11	23	51.1	10	11	US-09-799-250-392
12	23	51.1	10	11	US-09-799-250-397
13	23	51.1	10	11	US-09-799-250-492
14	23	51.1	10	11	US-09-799-250-502
15	23	51.1	10	11	US-09-799-250-575

16	23	51.1	10	11	US-09-799-250-610	Sequence 610, App
17	23	51.1	10	11	US-09-799-250-666	Sequence 666, App
18	23	51.1	10	11	US-09-799-250-699	Sequence 699, App
19	23	51.1	10	11	US-09-572-404B-3690	Sequence 3690, Ap
20	22	48.9	9	15	US-10-172-573-7	Sequence 7, Appli
21	22	48.9	10	11	US-09-572-404B-2178	Sequence 2178, Ap
22	21	46.7	9	12	US-10-357-175-78	Sequence 78, Appl
23	21	46.7	9	12	US-10-357-175-95	Sequence 95, Appl
24	21	46.7	9	12	US-10-357-175-146	Sequence 146, App
25	21	46.7	10	12	US-09-572-270A-469	Sequence 469, App
26	20	44.4	7	15	US-10-059-261-18	Sequence 18, Appl
27	20	44.4	7	15	US-10-059-261-126	Sequence 126, App
28	20	44.4	7	15	US-10-353-929-186	Sequence 186, App
29	20	44.4	10	9	US-09-854-204-52	Sequence 52, Appl
30	20	44.4	10	11	US-09-572-404B-1615	Sequence 1615, Ap
31	20	44.4	10	11	US-09-572-404B-1625	Sequence 1625, Ap
32	20	44.4	10	12	US-10-353-929-185	Sequence 185, App
33	19	42.2	5	9	US-09-096-749A-63	Sequence 63, Appl
34	19	42.2	5	11	US-09-903-412-63	Sequence 63, Appl
35	19	42.2	5	12	US-10-165-155-63	Sequence 63, Appl
36	19	42.2	5	12	US-10-190-162-63	Sequence 63, Appl
37	19	42.2	5	15	US-10-174-717A-63	Sequence 63, Appl
38	19	42.2	9	10	US-09-818-247-15	Sequence 15, Appl
39	19	42.2	9	10	US-09-792-480-17	Sequence 17, Appl
40	19	42.2	9	11	US-09-799-250-31	Sequence 31, Appl
41	19	42.2	9	11	US-09-799-250-158	Sequence 158, App
42	19	42.2	9	11	US-09-799-250-438	Sequence 438, App
43	19	42.2	9	11	US-09-799-250-524	Sequence 524, App
44	19	42.2	9	11	US-09-799-250-620	Sequence 620, App
45	19	42.2	9	11	US-09-865-348A-81	Sequence 81, Appl

ALIGNMENTS

RESULT 1  
US-09-799-250-40  
; Sequence 40, Application US/09799250  
; Publication No. US20030032087A1  
; GENERAL INFORMATION:  
; APPLICANT: Pia M. Challita-Eld  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Mary Paris  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: 121P1: A TISSUE SPECIFIC PROTEIN  
; FILE REFERENCE: 129.34US01  
; CURRENT APPLICATION NUMBER: US/09/799,250  
; CURRENT FILING DATE: 2001-02-08  
; NUMBER OF SEQ ID NOS: 719  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 40  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-799-250-40

Query Match 51.1%; Score 23; DB 11; Length 9;  
Best Local Similarity 57.1%; Pred. No. 5.2e+05;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 3 LAAQRR 9  
Db 1 LSAEKR 7

RESULT 2  
US-09-799-250-56  
; Sequence 56, Application US/09799250  
; Publication No. US20030032087A1

; GENERAL INFORMATION:  
 ; APPLICANT: Pia M. Challita-Eid  
 ; APPLICANT: Rene S. Hubert  
 ; APPLICANT: Steve Chappell Mitchell  
 ; APPLICANT: Arthur B. Raitano  
 ; APPLICANT: Mary Faris  
 ; APPLICANT: Daniel E.H. Afar  
 ; APPLICANT: Aya Jakobovits  
 ; TITLE OF INVENTION: 121PIF1: A TISSUE SPECIFIC PROTEIN  
 ; FILE REFERENCE: 129.34US01  
 ; CURRENT APPLICATION NUMBER: US/09/799,250  
 ; CURRENT FILING DATE: 2001-02-08  
 ; NUMBER OF SEQ ID NOS: 719  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 56  
 ; LENGTH: 9  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-799-250-56

Query Match 51.1%; Score 23; DB 11; Length 9;  
 Best Local Similarity 57.1%; Pred. No. 5.2e+05;  
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAAQERR 9  
 Db 3 LSAEKR 9

RESULT 3  
 US-09-799-250-121  
 ; Sequence 121, Application US/09799250  
 ; Publication No. US20030032087A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Pia M. Challita-Eid  
 ; APPLICANT: Rene S. Hubert  
 ; APPLICANT: Steve Chappell Mitchell  
 ; APPLICANT: Arthur B. Raitano  
 ; APPLICANT: Mary Faris  
 ; APPLICANT: Daniel E.H. Afar  
 ; APPLICANT: Aya Jakobovits  
 ; TITLE OF INVENTION: 121PIF1: A TISSUE SPECIFIC PROTEIN  
 ; FILE REFERENCE: 129.34US01  
 ; CURRENT APPLICATION NUMBER: US/09/799,250  
 ; CURRENT FILING DATE: 2001-02-08  
 ; NUMBER OF SEQ ID NOS: 719  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 121  
 ; LENGTH: 9  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-799-250-121

Query Match 51.1%; Score 23; DB 11; Length 9;  
 Best Local Similarity 57.1%; Pred. No. 5.2e+05;  
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAAQERR 9  
 Db 2 LSAEKR 8

RESULT 4  
 US-09-799-250-244  
 ; Sequence 244, Application US/09799250  
 ; Publication No. US20030032087A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Pia M. Challita-Eid  
 ; APPLICANT: Rene S. Hubert  
 ; APPLICANT: Steve Chappell Mitchell  
 ; APPLICANT: Arthur B. Raitano

; APPLICANT: Mary Faris  
 ; APPLICANT: Daniel E.H. Afar  
 ; APPLICANT: Aya Jakobovits  
 ; TITLE OF INVENTION: 121PIF1: A TISSUE SPECIFIC PROTEIN  
 ; FILE REFERENCE: 129.34US01  
 ; CURRENT APPLICATION NUMBER: US/09/799,250  
 ; CURRENT FILING DATE: 2001-02-08  
 ; NUMBER OF SEQ ID NOS: 719  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 244  
 ; LENGTH: 9  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-799-250-244

Query Match 51.1%; Score 23; DB 11; Length 9;  
 Best Local Similarity 57.1%; Pred. No. 5.2e+05;  
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAAQERR 9  
 Db 2 LSAEKR 8

RESULT 5  
 US-09-799-250-348  
 ; Sequence 348, Application US/09799250  
 ; Publication No. US20030032087A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Pia M. Challita-Eid  
 ; APPLICANT: Rene S. Hubert  
 ; APPLICANT: Steve Chappell Mitchell  
 ; APPLICANT: Arthur B. Raitano  
 ; APPLICANT: Mary Faris  
 ; APPLICANT: Daniel E.H. Afar  
 ; APPLICANT: Aya Jakobovits  
 ; TITLE OF INVENTION: 121PIF1: A TISSUE SPECIFIC PROTEIN  
 ; FILE REFERENCE: 129.34US01  
 ; CURRENT APPLICATION NUMBER: US/09/799,250  
 ; CURRENT FILING DATE: 2001-02-08  
 ; NUMBER OF SEQ ID NOS: 719  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 348  
 ; LENGTH: 9  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-799-250-348

Query Match 51.1%; Score 23; DB 11; Length 9;  
 Best Local Similarity 57.1%; Pred. No. 5.2e+05;  
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAAQERR 9  
 Db 3 LSAEKR 9

RESULT 6  
 US-09-799-250-560  
 ; Sequence 560, Application US/09799250  
 ; Publication No. US20030032087A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Pia M. Challita-Eid  
 ; APPLICANT: Rene S. Hubert  
 ; APPLICANT: Steve Chappell Mitchell  
 ; APPLICANT: Arthur B. Raitano  
 ; APPLICANT: Mary Faris  
 ; APPLICANT: Daniel E.H. Afar  
 ; APPLICANT: Aya Jakobovits  
 ; TITLE OF INVENTION: 121PIF1: A TISSUE SPECIFIC PROTEIN  
 ; FILE REFERENCE: 129.34US01



; FILE REFERENCE: 129.34US01  
; CURRENT APPLICATION NUMBER: US/09/799,250  
; CURRENT FILING DATE: 2001-02-08  
; NUMBER OF SEQ ID NOS: 719  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 560  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-799-250-560

Query Match 51.1%; Score 23; DB 11; Length 9;  
Best Local Similarity 57.1%; Pred. No. 5.2e+05;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAAQERR 9  
|:|:|:|  
DB 2 LSAEKR 8

RESULT 7  
US-09-799-250-93  
; Sequence 93, Application US/09799250  
; Publication No. US20030032087A1  
; GENERAL INFORMATION:  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Mary Faris  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: 121P1F1: A TISSUE SPECIFIC PROTEIN  
; FILE REFERENCE: 129.34US01  
; CURRENT APPLICATION NUMBER: US/09/799,250  
; CURRENT FILING DATE: 2001-02-08  
; NUMBER OF SEQ ID NOS: 719  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 93  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-799-250-93

Query Match 51.1%; Score 23; DB 11; Length 10;  
Best Local Similarity 57.1%; Pred. No. 2.4e+02;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAAQERR 9  
|:|:|:|  
DB 2 LSAEKR 8

RESULT 8  
US-09-799-250-178  
; Sequence 178, Application US/09799250  
; Publication No. US20030032087A1  
; GENERAL INFORMATION:  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Mary Faris  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: 121P1F1: A TISSUE SPECIFIC PROTEIN  
; FILE REFERENCE: 129.34US01  
; CURRENT APPLICATION NUMBER: US/09/799,250  
; CURRENT FILING DATE: 2001-02-08  
; NUMBER OF SEQ ID NOS: 719  
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 178  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-799-250-178

Query Match 51.1%; Score 23; DB 11; Length 10;  
Best Local Similarity 57.1%; Pred. No. 2.4e+02;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAAQERR 9  
|:|:|:|  
DB 3 LSAEKR 9

RESULT 9  
US-09-799-250-185  
; Sequence 185, Application US/09799250  
; Publication No. US20030032087A1  
; GENERAL INFORMATION:  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Mary Faris  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: 121P1F1: A TISSUE SPECIFIC PROTEIN  
; FILE REFERENCE: 129.34US01  
; CURRENT APPLICATION NUMBER: US/09/799,250  
; CURRENT FILING DATE: 2001-02-08  
; NUMBER OF SEQ ID NOS: 719  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 185  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-799-250-185

Query Match 51.1%; Score 23; DB 11; Length 10;  
Best Local Similarity 57.1%; Pred. No. 2.4e+02;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAAQERR 9  
|:|:|:|  
DB 1 LSAEKR 7

RESULT 10  
US-09-799-250-274  
; Sequence 274, Application US/09799250  
; Publication No. US20030032087A1  
; GENERAL INFORMATION:  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Mary Faris  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: 121P1F1: A TISSUE SPECIFIC PROTEIN  
; FILE REFERENCE: 129.34US01  
; CURRENT APPLICATION NUMBER: US/09/799,250  
; CURRENT FILING DATE: 2001-02-08  
; NUMBER OF SEQ ID NOS: 719  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 274  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-799-250-274

Query Match 51.1%; Score 23; DB 11; Length 10;  
Best Local Similarity 57.1%; Pred. No. 2.4e+02;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAAQERR 9  
|:|:|:|  
Db 2 LSAEKR 8

## RESULT 11

US-09-799-250-392  
; Sequence 392, Application US/09799250  
; Publication No. US20030032087A1  
; GENERAL INFORMATION:  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Mary Faris  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Ava Jakobovits  
; TITLE OF INVENTION: 121PFI: A TISSUE SPECIFIC PROTEIN  
; FILE REFERENCE: 129.34US01  
; CURRENT FILING DATE: 2001-02-08  
; NUMBER OF SEQ ID NOS: 719  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 392  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-799-250-392

Query Match 51.1%; Score 23; DB 11; Length 10;  
Best Local Similarity 57.1%; Pred. No. 2.4e+02;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAAQERR 9  
|:|:|:|  
Db 2 LSAEKR 8

## RESULT 12

US-09-799-250-397  
; Sequence 397, Application US/09799250  
; Publication No. US20030032087A1  
; GENERAL INFORMATION:  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Mary Faris  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Ava Jakobovits  
; TITLE OF INVENTION: 121PFI: A TISSUE SPECIFIC PROTEIN  
; FILE REFERENCE: 129.34US01  
; CURRENT FILING DATE: 2001-02-08  
; NUMBER OF SEQ ID NOS: 719  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 397  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-799-250-397

Query Match 51.1%; Score 23; DB 11; Length 10;  
Best Local Similarity 57.1%; Pred. No. 2.4e+02;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAAQERR 9  
|:|:|:|  
Db 4 LSAEKR 10

## RESULT 13

US-09-799-250-492  
; Sequence 492, Application US/09799250  
; Publication No. US20030032087A1  
; GENERAL INFORMATION:  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Mary Faris  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Ava Jakobovits  
; TITLE OF INVENTION: 121PFI: A TISSUE SPECIFIC PROTEIN  
; FILE REFERENCE: 129.34US01  
; CURRENT FILING DATE: 2001-02-08  
; NUMBER OF SEQ ID NOS: 719  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 492  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-799-250-492

Query Match 51.1%; Score 23; DB 11; Length 10;  
Best Local Similarity 57.1%; Pred. No. 2.4e+02;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAAQERR 9  
|:|:|:|  
Db 1 LSAEKR 7

## RESULT 14

US-09-799-250-502  
; Sequence 502, Application US/09799250  
; Publication No. US20030032087A1  
; GENERAL INFORMATION:  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Mary Faris  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Ava Jakobovits  
; TITLE OF INVENTION: 121PFI: A TISSUE SPECIFIC PROTEIN  
; FILE REFERENCE: 129.34US01  
; CURRENT FILING DATE: 2001-02-08  
; NUMBER OF SEQ ID NOS: 719  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 502  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-799-250-502

Query Match 51.1%; Score 23; DB 11; Length 10;  
Best Local Similarity 57.1%; Pred. No. 2.4e+02;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAAQERR 9  
|:|:|:|  
Db 3 LSAEKR 9

```
RESULT 15
US-09-799-250-575
; Sequence 575, Application US/09799250
; Publication No. US20030032087A1
; GENERAL INFORMATION:
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Rene S. Hubert
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Arthur B. Raitano
; APPLICANT: Mary Faris
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 121P1: A TISSUE SPECIFIC PROTEIN
; TITLE OF INVENTION: HIGHLY EXPRESSED IN VARIOUS CANCERS
; FILE REFERENCE: 129.34US01
; CURRENT APPLICATION NUMBER: US/09/799,250
; CURRENT FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 719
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 575
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-799-250-575

Query Match      51.1%; Score 23; DB 11; Length 10;
Best Local Similarity 57.1%; Pred. No. 2.4e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      3 LAAQERR 9
Db      1 LSAEEKR 7
```

Search completed: October 7, 2003, 14:43:30  
Job time : 42 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:27:56 ; Search time 276.5 Seconds  
(without alignments)  
32.908 Million cell updates/sec

Title: US-09-807-512-25

Perfect score: 45

Sequence: 1 AMLAAQERRV 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues

Total number of hits satisfying chosen parameters: 566550

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main:\*

1: /cgn2\_6/ptodata/1/paa/PCTUS\_COMB.pep.\*

2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*

3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep.\*

4: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*

5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep.\*

6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep.\*

7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep.\*

8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep.\*

9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep.\*

10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep.\*

11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep.\*

12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep.\*

13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep.\*

14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep.\*

15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep.\*

16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep.\*

17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep.\*

18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep.\*

19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep.\*

20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep.\*

21: /cgn2\_6/ptodata/1/paa/US097A\_COMB.pep.\*

22: /cgn2\_6/ptodata/1/paa/US097B\_COMB.pep.\*

23: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep.\*

24: /cgn2\_6/ptodata/1/paa/US099A\_COMB.pep.\*

25: /cgn2\_6/ptodata/1/paa/US099B\_COMB.pep.\*

26: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep.\*

27: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep.\*

28: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep.\*

29: /cgn2\_6/ptodata/1/paa/US103\_COMB.pep.\*

30: /cgn2\_6/ptodata/1/paa/US104\_COMB.pep.\*

31: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep.\*

32: /cgn2\_6/ptodata/1/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	45	100.0	10	23	US-09-807-512-25
					Sequence 25, Appl

2	41	91.1	9	23	US-09-807-512-26	Sequence 26, Appl
3	36	80.0	9	19	US-09-529-206A-123	Sequence 123, App
4	36	80.0	9	19	US-09-529-206B-123	Sequence 123, App
5	36	80.0	9	19	US-09-529-206D-123	Sequence 123, App
6	36	80.0	10	17	US-09-336-091-45	Sequence 45, Appl
7	36	80.0	10	19	US-09-529-206-47	Sequence 47, Appl
8	36	80.0	10	19	US-09-529-206A-47	Sequence 47, Appl
9	36	80.0	10	19	US-09-529-206B-47	Sequence 47, Appl
10	36	80.0	10	19	US-09-529-206D-47	Sequence 47, Appl
11	32	71.1	9	19	US-09-529-206-46	Sequence 46, Appl
12	32	71.1	9	19	US-09-529-206A-46	Sequence 46, Appl
13	32	71.1	9	19	US-09-529-206B-46	Sequence 46, Appl
14	32	71.1	9	19	US-09-529-206D-46	Sequence 46, Appl
15	30.5	67.8	10	19	US-09-529-206A-124	Sequence 124, App
16	30.5	67.8	10	19	US-09-529-206B-124	Sequence 124, App
17	30.5	67.8	10	19	US-09-529-206D-124	Sequence 124, App
18	25	55.6	9	5	US-08-144-779-27	Sequence 27, Appl
19	24	53.3	9	27	US-10-119-536A-181	Sequence 181, App
20	23	51.1	9	22	US-09-799-250-40	Sequence 40, Appl
21	23	51.1	9	22	US-09-799-250-56	Sequence 56, Appl
22	23	51.1	9	22	US-09-799-250-121	Sequence 121, App
23	23	51.1	9	22	US-09-799-250-244	Sequence 244, App
24	23	51.1	9	22	US-09-799-250-348	Sequence 348, App
25	23	51.1	9	22	US-09-799-250-560	Sequence 560, App
26	23	51.1	9	22	US-09-799-250A-40	Sequence 40, Appl
27	23	51.1	9	22	US-09-799-250A-56	Sequence 56, Appl
28	23	51.1	9	22	US-09-799-250A-121	Sequence 121, App
29	23	51.1	9	22	US-09-799-250A-244	Sequence 244, App
30	23	51.1	9	22	US-09-799-250A-348	Sequence 348, App
31	23	51.1	9	22	US-09-799-250A-560	Sequence 560, App
32	23	51.1	10	19	US-09-572-404B-3690	Sequence 3690, Ap
33	23	51.1	10	22	US-09-799-250-93	Sequence 93, Appl
34	23	51.1	10	22	US-09-799-250-178	Sequence 178, App
35	23	51.1	10	22	US-09-799-250-185	Sequence 185, App
36	23	51.1	10	22	US-09-799-250-274	Sequence 274, App
37	23	51.1	10	22	US-09-799-250-392	Sequence 392, App
38	23	51.1	10	22	US-09-799-250-397	Sequence 397, App
39	23	51.1	10	22	US-09-799-250-492	Sequence 492, App
40	23	51.1	10	22	US-09-799-250-502	Sequence 502, App
41	23	51.1	10	22	US-09-799-250-575	Sequence 575, App
42	23	51.1	10	22	US-09-799-250-610	Sequence 610, App
43	23	51.1	10	22	US-09-799-250-666	Sequence 666, App
44	23	51.1	10	22	US-09-799-250-699	Sequence 699, App
45	23	51.1	10	22	US-09-799-250A-93	Sequence 93, Appl

ALIGNMENTS

RESULT 1

US-09-807-512-25

; Sequence 25, Application US/09807512

; GENERAL INFORMATION:

; APPLICANT: Schrier, Peter I.

; APPLICANT: Aarnoudse, Corlien

; APPLICANT: Heider, Karl-Heinz

; APPLICANT: Klade, Christoph

; TITLE OF INVENTION: Camel, An Alternative Translation Product of the Tumor

; FILE OF INVENTION: Antigen-Lage 1

; FILE REFERENCE: 0652.2200000

; CURRENT APPLICATION NUMBER: US/09/807,512

; CURRENT FILING DATE: 2001-04-16

; PRIOR APPLICATION NUMBER: PCT/EP99/07832

; PRIOR FILING DATE: 1999-10-15

; PRIOR APPLICATION NUMBER: EP 98119583.7

; PRIOR FILING DATE: 1998-10-16

; NUMBER OF SEQ ID NOS: 28

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 25

; LENGTH: 10

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-807-512-25

```
Query Match      100.0%; Score 45; DB 23; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AMLAAQERRV 10
        |||||
Db       1 AMLAAQERRV 10

RESULT 2
US-09-807-512-26
; Sequence 26, Application US/09807512
; GENERAL INFORMATION:
; APPLICANT: Schrier, Peter I.
; APPLICANT: Aarnoudse, Corlien
; APPLICANT: Heider, Karl-Heinz
; APPLICANT: Klade, Christoph
; TITLE OF INVENTION: Camel, An Alternative Translation Product of the Tumor
; TITLE OF INVENTION: Antigen-Lage 1
; FILE REFERENCE: 0652.2200000
; CURRENT APPLICATION NUMBER: US/09/807,512
; CURRENT FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: PCT/EP99/07832
; PRIOR FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: EP 98119583.7
; PRIOR FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-807-512-26

Query Match      91.1%; Score 41; DB 23; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.3e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 MAAQERRV 10
        |||||
Db       1 MAAQERRV 9

RESULT 3
US-09-529-206A-123
; Sequence 123, Application US/09529206A
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; TITLE OF INVENTION: Encoding Same
; FILE REFERENCE: 20264269US1
; CURRENT APPLICATION NUMBER: US/09/529,206A
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: PCT/US98/19609
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 123
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206A-123

Query Match      80.0%; Score 36; DB 19; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.3e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AMLAAQER 8
        |||||
Db       2 AMLAAQER 9

RESULT 4
US-09-529-206B-123
; Sequence 123, Application US/09529206B
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; TITLE OF INVENTION: Encoding Same
; FILE REFERENCE: 20264269US1
; CURRENT APPLICATION NUMBER: US/09/529,206B
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: PCT/US98/19609
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 123
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206B-123

Query Match      80.0%; Score 36; DB 19; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.3e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AMLAAQER 8
        |||||
Db       2 AMLAAQER 9

RESULT 5
US-09-529-206D-123
; Sequence 123, Application US/09529206D
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; TITLE OF INVENTION: Encoding Same
; FILE REFERENCE: 218791
; CURRENT APPLICATION NUMBER: US/09/529,206D
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: PCT/US98/19609
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 123
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206D-123

Query Match      80.0%; Score 36; DB 19; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.3e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AMLAAQER 8
        |||||
Db       2 AMLAAQER 9

RESULT 6
US-09-336-091-45
; Sequence 45, Application US/09336091
; GENERAL INFORMATION:
; APPLICANT: Van Snick, Jacques
; APPLICANT: Leth, Bernard
```

; APPLICANT: Chauv. Pascal  
; APPLICANT: Boon-Falleur, Thierry  
; APPLICANT: van der Bruggen, Pierre  
; TITLE OF INVENTION: MAGE-A1 PEPTIDES PRESENTED BY HLA CLASS  
; FILE REFERENCE: L0461/7063  
; CURRENT APPLICATION NUMBER: US/09/336,091  
; CURRENT FILING DATE: 1999-06-18  
; NUMBER OF SEQ ID NOS: 72  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 45  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-336-091-45

Query Match 80.0%; Score 36; DB 17; Length 10;  
Best Local Similarity 100.0%; Pred. No. 2.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAAQERRV 10  
| | | | | | | |  
Db 1 LAAQERRV 8

RESULT 7  
US-09-529-206-47  
; Sequence 47, Application US/09529206  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America as  
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene  
; FILE REFERENCE: 20264269PC  
; CURRENT APPLICATION NUMBER: US/09/529,206  
; PRIOR FILING DATE: 2000-06-13  
; PRIOR APPLICATION NUMBER: US60/061,428  
; PRIOR FILING DATE: 1997-10-08  
; NUMBER OF SEQ ID NOS: 106  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 47  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-206-47

Query Match 80.0%; Score 36; DB 19; Length 10;  
Best Local Similarity 100.0%; Pred. No. 2.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAAQERRV 10  
| | | | | | | |  
Db 1 LAAQERRV 8

RESULT 8  
US-09-529-206A-47  
; Sequence 47, Application US/09529206A  
; GENERAL INFORMATION:  
; APPLICANT: Wang Rong, Fu  
; APPLICANT: Rosenberg, Steven  
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene  
; FILE REFERENCE: 20264269US1  
; CURRENT APPLICATION NUMBER: US/09/529,206A  
; PRIOR FILING DATE: 2000-06-13  
; PRIOR APPLICATION NUMBER: PCT/US98/19609  
; PRIOR FILING DATE: 1998-09-21  
; PRIOR APPLICATION NUMBER: US60/061,428  
; PRIOR FILING DATE: 1997-10-08  
; NUMBER OF SEQ ID NOS: 127  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 47  
; LENGTH: 10

; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-206A-47

Query Match 80.0%; Score 36; DB 19; Length 10;  
Best Local Similarity 100.0%; Pred. No. 2.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAAQERRV 10  
| | | | | | | |  
Db 1 LAAQERRV 8

RESULT 9  
US-09-529-206B-47  
; Sequence 47, Application US/09529206B  
; GENERAL INFORMATION:  
; APPLICANT: Wang Rong, Fu  
; APPLICANT: Rosenberg, Steven  
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene  
; FILE REFERENCE: 20264269US1  
; CURRENT APPLICATION NUMBER: US/09/529,206B  
; CURRENT FILING DATE: 2000-06-13  
; PRIOR APPLICATION NUMBER: PCT/US98/19609  
; PRIOR FILING DATE: 1998-09-21  
; PRIOR APPLICATION NUMBER: US60/061,428  
; PRIOR FILING DATE: 1997-10-08  
; NUMBER OF SEQ ID NOS: 127  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 47  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-206B-47

Query Match 80.0%; Score 36; DB 19; Length 10;  
Best Local Similarity 100.0%; Pred. No. 2.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAAQERRV 10  
| | | | | | | |  
Db 1 LAAQERRV 8

RESULT 10  
US-09-529-206D-47  
; Sequence 47, Application US/09529206D  
; GENERAL INFORMATION:  
; APPLICANT: Wang Rong, Fu  
; APPLICANT: Rosenberg, Steven  
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene  
; FILE REFERENCE: 218791  
; CURRENT APPLICATION NUMBER: US/09/529,206D  
; CURRENT FILING DATE: 2000-06-13  
; PRIOR APPLICATION NUMBER: PCT/US98/19609  
; PRIOR FILING DATE: 1998-09-21  
; PRIOR APPLICATION NUMBER: US60/061,428  
; PRIOR FILING DATE: 1997-10-08  
; NUMBER OF SEQ ID NOS: 127  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 47  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-206D-47

Query Match 80.0%; Score 36; DB 19; Length 10;  
Best Local Similarity 100.0%; Pred. No. 2.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAAQERRV 10

```

Db          1 LAAQERRV 8
|||||||
FILE REFERENCE: 20264269US1
; CURRENT APPLICATION NUMBER: US/09/529,206B
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: PCT/US98/19609
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206B-46

Query Match      71.1%; Score 32; DB 19; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.3e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 AAQERRV 10
        |||||||
DB       1 AAQERRV 7

RESULT 14
US-09-529-206D-46
; Sequence 46, Application US/09529206D
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; TITLE OF INVENTION: Encoding Same
; FILE REFERENCE: 218791
; CURRENT APPLICATION NUMBER: US/09/529,206D
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: PCT/US98/19609
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206D-46

Query Match      71.1%; Score 32; DB 19; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.3e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 AAQERRV 10
        |||||||
DB       1 AAQERRV 7

RESULT 15
US-09-529-206A-124
; Sequence 124, Application US/09529206A
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; TITLE OF INVENTION: Encoding Same
; FILE REFERENCE: 20264269US1
; CURRENT APPLICATION NUMBER: US/09/529,206A
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: PCT/US98/19609
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1

```

---

```

Db          1 LAAQERRV 8
|||||||
FILE REFERENCE: 20264269US1
; CURRENT APPLICATION NUMBER: US/09/529,206B
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: PCT/US98/19609
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206B-46

Query Match      71.1%; Score 32; DB 19; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.3e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 AAQERRV 10
        |||||||
DB       1 AAQERRV 7

RESULT 12
US-09-529-206A-46
; Sequence 46, Application US/09529206A
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; TITLE OF INVENTION: Encoding Same
; FILE REFERENCE: 20264269US1
; CURRENT APPLICATION NUMBER: US/09/529,206A
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: PCT/US98/19609
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206A-46

Query Match      71.1%; Score 32; DB 19; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.3e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 AAQERRV 10
        |||||||
DB       1 AAQERRV 7

RESULT 13
US-09-529-206B-46
; Sequence 46, Application US/09529206B
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; TITLE OF INVENTION: Encoding Same

```



; SEQ ID NO 124  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-206A-124

Query Match 67.8%; Score 30.5; DB 19; Length 10;  
Best Local Similarity 90.0%; Pred. NO. 39;  
Matches 9; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
Qy 1 AMLAA-OERR 9  
      |||||  
Db 1 AMLAARGERR 10

Search completed: October 7, 2003, 14:41:24  
Job time : 276.5 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:28:26 ; Search time 14.5 Seconds  
(without alignments)  
20.087 Million cell updates/sec

Title: US-09-807-512-25

Perfect score: 45

Sequence: 1 AMLAAQRRV 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 136899 seqs, 29126274 residues

Total number of hits satisfying chosen parameters: 23559

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_New:\*

1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	23	51.1	9	5	US-09-799-250B-40
2	23	51.1	9	5	US-09-799-250B-56
3	23	51.1	9	5	US-09-799-250B-121
4	23	51.1	9	5	US-09-799-250B-244
5	23	51.1	9	5	US-09-799-250B-348
6	23	51.1	9	5	US-09-799-250B-560
7	23	51.1	10	5	US-09-799-250B-93
8	23	51.1	10	5	US-09-799-250B-178
9	23	51.1	10	5	US-09-799-250B-185
10	23	51.1	10	5	US-09-799-250B-392
11	23	51.1	10	5	US-09-799-250B-372
12	23	51.1	10	5	US-09-799-250B-397
13	23	51.1	10	5	US-09-799-250B-492
14	23	51.1	10	5	US-09-799-250B-502
15	23	51.1	10	5	US-09-799-250B-575
16	23	51.1	10	5	US-09-799-250B-610
17	23	51.1	10	5	US-09-799-250B-666
18	23	51.1	10	5	US-09-799-250B-699
19	21	46.7	7	5	US-09-627-600C-43
20	21	46.7	7	5	US-09-627-600C-43
21	21	46.7	8	5	US-09-627-600C-93
22	21	46.7	8	5	US-09-627-600C-93
23	20	44.4	5	5	US-09-627-600C-124
24	20	44.4	5	5	US-09-627-600C-124
25	20	44.4	8	5	US-09-390-061D-251
26	20	44.4	8	5	US-09-390-061D-2053

27 20 44.4 8 5 US-09-390-061D-2240 Sequence 2240, Ap  
28 20 44.4 8 5 US-09-390-061D-2256 Sequence 2256, Ap  
29 20 44.4 8 5 US-09-390-061D-2883 Sequence 2883, Ap  
30 20 44.4 8 5 US-09-390-061D-2898 Sequence 2898, Ap  
31 20 44.4 8 5 US-09-627-600C-103 Sequence 103, App  
32 20 44.4 8 5 US-09-627-600D-103 Sequence 103, App  
33 20 44.4 9 1 PCT-US02-27760-11 Sequence 32, Appl  
34 20 44.4 9 1 PCT-US02-27760-32 Sequence 32, Appl  
35 20 44.4 9 5 US-09-390-061D-616 Sequence 616, App  
36 20 44.4 9 5 US-09-390-061D-698 Sequence 698, App  
37 20 44.4 9 5 US-09-390-061D-2054 Sequence 2054, Ap  
38 20 44.4 9 5 US-09-390-061D-2223 Sequence 2223, Ap  
39 20 44.4 9 5 US-09-390-061D-2241 Sequence 2241, Ap  
40 20 44.4 9 5 US-09-390-061D-2257 Sequence 2257, Ap  
41 20 44.4 9 5 US-09-390-061D-2700 Sequence 2700, Ap  
42 20 44.4 9 5 US-09-390-061D-2871 Sequence 2871, Ap  
43 20 44.4 9 5 US-09-390-061D-3515 Sequence 3515, Ap  
44 20 44.4 10 1 PCT-US02-27760-80 Sequence 80, Appl  
45 20 44.4 10 1 PCT-US02-27760-400 Sequence 400, App

#### ALIGNMENTS

RESULT 1  
US-09-799-250B-40  
; Sequence 40, Application US/09799250B  
; GENERAL INFORMATION:  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Mary Faris  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: 121P1: A TISSUE SPECIFIC PROTEIN  
; FILE REFERENCE: 129.34US01 (511582003400)  
; CURRENT APPLICATION NUMBER: US/09/799,250B  
; CURRENT FILING DATE: 2003-07-14  
; NUMBER OF SEQ ID NOS: 721  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 40  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-799-250B-40

Query Match 51.1%; Score 23; DB 5; Length 9;  
Best Local Similarity 57.1%; Pred. No. 1.2e+05;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 3 LAAQRR 9  
I::I::I  
Db 1 LSAEKR 7

RESULT 2  
US-09-799-250B-56  
; Sequence 56, Application US/09799250B  
; GENERAL INFORMATION:  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Mary Faris  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: 121P1: A TISSUE SPECIFIC PROTEIN  
; FILE REFERENCE: 129.34US01 (511582003400)  
; CURRENT APPLICATION NUMBER: US/09/799,250B  
; CURRENT FILING DATE: 2003-07-14

```
; NUMBER OF SEQ ID NOS: 721
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 56
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-799-250B-56

Query Match      51.1%; Score 23; DB 5; Length 9;
Best Local Similarity 57.1%; Pred. No. 1.2e+05;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      3 LAAQERR 9
Db      2 LSAEKKR 8

RESULT 3
US-09-799-250B-121
; Sequence 121, Application US/09799250B
; GENERAL INFORMATION:
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Rene S. Hubert
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Arthur B. Raitano
; APPLICANT: Mary Faris
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 121P1F1: A TISSUE SPECIFIC PROTEIN
; FILE REFERENCE: 129.34US01 (511582003400)
; CURRENT APPLICATION NUMBER: US/09/799,250B
; CURRENT FILING DATE: 2003-07-14
; NUMBER OF SEQ ID NOS: 721
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 121
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-799-250B-121

Query Match      51.1%; Score 23; DB 5; Length 9;
Best Local Similarity 57.1%; Pred. No. 1.2e+05;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      3 LAAQERR 9
Db      2 LSAEKKR 8

RESULT 4
US-09-799-250B-244
; Sequence 244, Application US/09799250B
; GENERAL INFORMATION:
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Rene S. Hubert
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Arthur B. Raitano
; APPLICANT: Mary Faris
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 121P1F1: A TISSUE SPECIFIC PROTEIN
; FILE REFERENCE: 129.34US01 (511582003400)
; CURRENT APPLICATION NUMBER: US/09/799,250B
; CURRENT FILING DATE: 2003-07-14
; NUMBER OF SEQ ID NOS: 721
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 244
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-799-250B-244

Query Match      51.1%; Score 23; DB 5; Length 9;
Best Local Similarity 57.1%; Pred. No. 1.2e+05;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      3 LAAQERR 9
Db      2 LSAEKKR 8

RESULT 5
US-09-799-250B-348
; Sequence 348, Application US/09799250B
; GENERAL INFORMATION:
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Rene S. Hubert
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Arthur B. Raitano
; APPLICANT: Mary Faris
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 121P1F1: A TISSUE SPECIFIC PROTEIN
; FILE REFERENCE: 129.34US01 (511582003400)
; CURRENT APPLICATION NUMBER: US/09/799,250B
; CURRENT FILING DATE: 2003-07-14
; NUMBER OF SEQ ID NOS: 721
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 348
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-799-250B-348

Query Match      51.1%; Score 23; DB 5; Length 9;
Best Local Similarity 57.1%; Pred. No. 1.2e+05;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      3 LAAQERR 9
Db      2 LSAEKKR 8

RESULT 6
US-09-799-250B-560
; Sequence 560, Application US/09799250B
; GENERAL INFORMATION:
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Rene S. Hubert
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Arthur B. Raitano
; APPLICANT: Mary Faris
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 121P1F1: A TISSUE SPECIFIC PROTEIN
; FILE REFERENCE: 129.34US01 (511582003400)
; CURRENT APPLICATION NUMBER: US/09/799,250B
; CURRENT FILING DATE: 2003-07-14
; NUMBER OF SEQ ID NOS: 721
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 560
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-799-250B-560

Query Match      51.1%; Score 23; DB 5; Length 9;
Best Local Similarity 57.1%; Pred. No. 1.2e+05;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      3 LAAQERR 9
Db      2 LSAEKKR 8

RESULT 7
US-09-799-250B-721
; Sequence 721, Application US/09799250B
; GENERAL INFORMATION:
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Rene S. Hubert
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Arthur B. Raitano
; APPLICANT: Mary Faris
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 121P1F1: A TISSUE SPECIFIC PROTEIN
; FILE REFERENCE: 129.34US01 (511582003400)
; CURRENT APPLICATION NUMBER: US/09/799,250B
; CURRENT FILING DATE: 2003-07-14
; NUMBER OF SEQ ID NOS: 721
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 721
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-799-250B-721

Query Match      51.1%; Score 23; DB 5; Length 9;
Best Local Similarity 57.1%; Pred. No. 1.2e+05;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      3 LAAQERR 9
Db      2 LSAEKKR 8
```

Db 2 LSAEKR 8

## RESULT 7

US-09-799-250B-93

; Sequence 93, Application US/09799250B

; GENERAL INFORMATION:

; APPLICANT: Pia M. Challita-Eid

; APPLICANT: Rene S. Hubert

; APPLICANT: Steve Chappell Mitchell

; APPLICANT: Arthur B. Raitano

; APPLICANT: Mary Faris

; APPLICANT: Daniel E.H. Afar

; APPLICANT: Aya Jakobovits

; TITLE OF INVENTION: 12IP1: A TISSUE SPECIFIC PROTEIN

; FILE OF INVENTION: HIGHLY EXPRESSED IN VARIOUS CANCERS

; FILE REFERENCE: 129.34US01 (511582003400)

; CURRENT APPLICATION NUMBER: US/09799,250B

; CURRENT FILING DATE: 2003-07-14

; NUMBER OF SEQ ID NOS: 721

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 93

; LENGTH: 10

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-799-250B-93

Query Match

Best Local Similarity 51.1%; Score 23; DB 5; Length 10;

Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAAQERR 9

I::I::I

Db 2 LSAEKR 8

## RESULT 8

US-09-799-250B-178

; Sequence 178, Application US/09799250B

; GENERAL INFORMATION:

; APPLICANT: Pia M. Challita-Eid

; APPLICANT: Rene S. Hubert

; APPLICANT: Steve Chappell Mitchell

; APPLICANT: Arthur B. Raitano

; APPLICANT: Mary Faris

; APPLICANT: Daniel E.H. Afar

; APPLICANT: Aya Jakobovits

; TITLE OF INVENTION: 12IP1: A TISSUE SPECIFIC PROTEIN

; FILE OF INVENTION: HIGHLY EXPRESSED IN VARIOUS CANCERS

; FILE REFERENCE: 129.34US01 (511582003400)

; CURRENT APPLICATION NUMBER: US/09799,250B

; CURRENT FILING DATE: 2003-07-14

; NUMBER OF SEQ ID NOS: 721

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 178

; LENGTH: 10

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-799-250B-178

Query Match

Best Local Similarity 51.1%; Score 23; DB 5; Length 10;

Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAAQERR 9

I::I::I

Db 3 LSAEKR 9

## RESULT 9

US-09-799-250B-185

; Sequence 185, Application US/09799250B

; GENERAL INFORMATION:

; APPLICANT: Pia M. Challita-Eid

; APPLICANT: Rene S. Hubert

; APPLICANT: Steve Chappell Mitchell

; APPLICANT: Arthur B. Raitano

; APPLICANT: Mary Faris

; APPLICANT: Daniel E.H. Afar

; APPLICANT: Aya Jakobovits

; APPLICANT: Pia M. Challita-Eid

; APPLICANT: Rene S. Hubert

; APPLICANT: Steve Chappell Mitchell

; APPLICANT: Arthur B. Raitano

; APPLICANT: Mary Faris

; APPLICANT: Daniel E.H. Afar

; APPLICANT: Aya Jakobovits

; TITLE OF INVENTION: 12IP1: A TISSUE SPECIFIC PROTEIN

; FILE OF INVENTION: HIGHLY EXPRESSED IN VARIOUS CANCERS

; FILE REFERENCE: 129.34US01 (511582003400)

; CURRENT APPLICATION NUMBER: US/09799,250B

; CURRENT FILING DATE: 2003-07-14

; NUMBER OF SEQ ID NOS: 721

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 185

; LENGTH: 10

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-799-250B-185

Query Match

Best Local Similarity 51.1%; Score 23; DB 5; Length 10;

Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAAQERR 9

I::I::I

Db 1 LSAEKR 7

## RESULT 10

US-09-799-250B-274

; Sequence 274, Application US/09799250B

; GENERAL INFORMATION:

; APPLICANT: Pia M. Challita-Eid

; APPLICANT: Rene S. Hubert

; APPLICANT: Steve Chappell Mitchell

; APPLICANT: Arthur B. Raitano

; APPLICANT: Mary Faris

; APPLICANT: Daniel E.H. Afar

; APPLICANT: Aya Jakobovits

; TITLE OF INVENTION: 12IP1: A TISSUE SPECIFIC PROTEIN

; FILE OF INVENTION: HIGHLY EXPRESSED IN VARIOUS CANCERS

; FILE REFERENCE: 129.34US01 (511582003400)

; CURRENT APPLICATION NUMBER: US/09799,250B

; CURRENT FILING DATE: 2003-07-14

; NUMBER OF SEQ ID NOS: 721

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 274

; LENGTH: 10

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-799-250B-274

Query Match

Best Local Similarity 51.1%; Score 23; DB 5; Length 10;

Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAAQERR 9

I::I::I

Db 2 LSAEKR 8

## RESULT 11

US-09-799-250B-392

; Sequence 392, Application US/09799250B

; GENERAL INFORMATION:

; APPLICANT: Pia M. Challita-Eid

; APPLICANT: Rene S. Hubert

; APPLICANT: Steve Chappell Mitchell

; APPLICANT: Arthur B. Raitano

; APPLICANT: Mary Faris

; APPLICANT: Daniel E.H. Afar

; APPLICANT: Aya Jakobovits

US-09-799-250B-392

; TITLE OF INVENTION: 121P1F1: A TISSUE SPECIFIC PROTEIN  
; TITLE OF INVENTION: HIGHLY EXPRESSED IN VARIOUS CANCERS  
; FILE REFERENCE: 129.34US01 (511582003400)  
; CURRENT APPLICATION NUMBER: US/09/799,250B  
; CURRENT FILING DATE: 2003-07-14  
; NUMBER OF SEQ ID NOS: 721  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 392  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-799-250B-392

Query Match 51.1%; Score 23; DB 5; Length 10;  
Best Local Similarity 57.1%; Pred. No. 1.4e+02;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAAQERR 9

Db 2 LSAEKR 8

#### RESULT 12

US-09-799-250B-397

; Sequence 397, Application US/09799250B  
; GENERAL INFORMATION:  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Mary Faris  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: 121P1F1: A TISSUE SPECIFIC PROTEIN  
; TITLE OF INVENTION: HIGHLY EXPRESSED IN VARIOUS CANCERS  
; FILE REFERENCE: 129.34US01 (511582003400)  
; CURRENT APPLICATION NUMBER: US/09/799,250B  
; CURRENT FILING DATE: 2003-07-14  
; NUMBER OF SEQ ID NOS: 721  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 397  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-799-250B-397

Query Match 51.1%; Score 23; DB 5; Length 10;  
Best Local Similarity 57.1%; Pred. No. 1.4e+02;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAAQERR 9

Db 4 LSAEKR 10

#### RESULT 13

US-09-799-250B-492

; Sequence 492, Application US/09799250B  
; GENERAL INFORMATION:  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Mary Faris  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: 121P1F1: A TISSUE SPECIFIC PROTEIN  
; TITLE OF INVENTION: HIGHLY EXPRESSED IN VARIOUS CANCERS  
; FILE REFERENCE: 129.34US01 (511582003400)  
; CURRENT APPLICATION NUMBER: US/09/799,250B  
; CURRENT FILING DATE: 2003-07-14  
; NUMBER OF SEQ ID NOS: 721  
; SOFTWARE: FastSeq for Windows Version 4.0

Query Match 51.1%; Score 23; DB 5; Length 10;

; SEQ ID NO 492  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-799-250B-492

Query Match 51.1%; Score 23; DB 5; Length 10;  
Best Local Similarity 57.1%; Pred. No. 1.4e+02;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAAQERR 9

Db 1 LSAEKR 7

#### RESULT 14

US-09-799-250B-502

; Sequence 502, Application US/09799250B  
; GENERAL INFORMATION:  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Mary Faris  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: 121P1F1: A TISSUE SPECIFIC PROTEIN  
; TITLE OF INVENTION: HIGHLY EXPRESSED IN VARIOUS CANCERS  
; FILE REFERENCE: 129.34US01 (511582003400)  
; CURRENT APPLICATION NUMBER: US/09/799,250B  
; CURRENT FILING DATE: 2003-07-14  
; NUMBER OF SEQ ID NOS: 721  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 502  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-799-250B-502

Query Match 51.1%; Score 23; DB 5; Length 10;  
Best Local Similarity 57.1%; Pred. No. 1.4e+02;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAAQERR 9

Db 3 LSAEKR 9

#### RESULT 15

US-09-799-250B-575

; Sequence 575, Application US/09799250B  
; GENERAL INFORMATION:  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Mary Faris  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: 121P1F1: A TISSUE SPECIFIC PROTEIN  
; TITLE OF INVENTION: HIGHLY EXPRESSED IN VARIOUS CANCERS  
; FILE REFERENCE: 129.34US01 (511582003400)  
; CURRENT APPLICATION NUMBER: US/09/799,250B  
; CURRENT FILING DATE: 2003-07-14  
; NUMBER OF SEQ ID NOS: 721  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 575  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-799-250B-575

Query Match

Best Local Similarity 57.1%; Pred. No. 1.4e+02;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
Oy 3 LAAQERR 9  
|:|:|:|  
Db 1 LSAREKR 7

Search completed: October 7, 2003, 14:42:01  
Job time : 15.5 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:31:21 ; Search time 57.5 Seconds  
(without alignments)  
24.844 Million cell updates/sec

Title: US-09-807-512-26  
Perfect score: 41  
Sequence: 1 MAAQERRV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues  
Total number of hits satisfying chosen parameters: 179625

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_19Jun03:\*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*

2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*

3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*

4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*

5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*

6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*

7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*

8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*

9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*

10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*

11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*

12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*

13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*

14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*

15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*

16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*

17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*

18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*

19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*

20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*

21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*

22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*

23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	100.0	9	21	AA70859
2	32	78.0	9	20	AA705981
3	25	61.0	9	21	AA809858
4	25	61.0	9	23	AAE20695
5	24	58.5	9	24	ABP75370
6	22	53.7	7	21	AB07254
7	22	53.7	7	23	AAE20692
8	22	53.7	8	14	AA36590
9	22	53.7	8	21	AA807255

10	22	53.7	8	23	AAE20693
11	22	53.7	9	21	AA809857
12	22	53.7	9	22	AA869558
13	22	53.7	9	23	AAE20694
14	21	51.2	6	20	AAE23550
15	21	51.2	7	22	AA806661
16	21	51.2	7	23	AA85848
17	21	51.2	8	22	AA80698
18	21	51.2	9	20	AAW97915
19	21	51.2	9	21	AA832294
20	21	51.2	9	21	AA832311
21	21	51.2	9	21	AA832362
22	21	51.2	9	21	AA832362
23	21	51.2	9	22	AAU68976
24	21	51.2	9	22	AAU68993
25	21	51.2	9	22	AAU69044
26	21	51.2	9	22	AA898549
27	20	48.8	5	22	AA80741
28	20	48.8	6	22	AA866768
29	20	48.8	7	16	AA877508
30	20	48.8	7	18	AAW22453
31	20	48.8	7	19	AAW71011
32	20	48.8	7	24	ABG71797
33	20	48.8	8	20	AAW74082
34	20	48.8	8	21	AA821665
35	20	48.8	8	21	AA809862
36	20	48.8	8	22	AA80709
37	20	48.8	8	23	ABP61615
38	20	48.8	8	23	AAE20699
39	20	48.8	9	17	AA893129
40	20	48.8	9	20	AAW84253
41	20	48.8	9	20	AAW84219
42	20	48.8	9	23	AB897125
43	20	48.8	9	24	ABR07493
44	20	48.8	9	24	ABR07505
45	20	48.8	9	24	ABR07558

ALIGNMENTS

RESULT 1

AA70859  
ID AA70859 standard; peptide; 9 AA.  
XX  
AC AA70859;  
XX  
DT 31-JUL-2000 (first entry)  
DE CAMEL17 immunogenic peptide of human CAMEL protein.

CAMEL; CTL-recognised Antigen on MELanoma; cytotoxic T lymphocyte; CTL;  
tumour-associated antigen; LAGE-1; NY-ESO-1; anticancer; melanoma;  
human; cancer; immunotherapy; immunogenic peptide; immune response.

XX Homo sapiens.

XX WO200023584-A1.

XX 27-APR-2000.

XX 15-OCT-1999; 99WO-EP07832.

XX 16-OCT-1998; 98EP-0119583.

XX (BOEH) BOEHRINGER INGELHEIM INT GMBH.

XX (UYHO-) UNIV HOSPITAL LEIDEN.

XX Schrier PI, Aarnoudse CA, Heider K, Klade C;

XX WPI; 2000-339685/29.

XX Tumor-associated antigen useful for cancer immunotherapy is encoded by

the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA -

Claim 8; Page 34; 73pp; English.

The present sequence is an immunogenic peptide CAMEL 17, of the human tumour-associated antigen CAMEL (Cytotoxic T lymphocytes (CTL)-recognised Antigen on Melanoma). This peptide has the potential to bind to HLA-A2 and corresponds to residues 17-25 of the CAMEL protein. The CAMEL protein is encoded by the LAGE-1 gene, a tumour-specific antigen. It is different from the LAGE-1 protein, since it is translated from a different open reading frame (ORF-1). It shows strong homology with NY-ESO-1, a melanoma specific tumour antigen. The tumour-associated antigen displayed on melanoma cells is recognised by cytotoxic T lymphocytes. CAMEL is expressed in tumour cell lines, tumour tissues (e.g. breast and lung) and in restricted number of healthy tissues. This sequence has anticancer activity. CAMEL tumour antigen and immunogenic peptides derived from it are useful for cancer immunotherapy. They have the potential to induce an immune response, by eliciting a CTL response. The DNA molecule is used to construct recombinant or fusion proteins.

Query Match 100.0%; Score 41; DB 21; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAAQRRV 9  
 |||||  
 Db 1 MLAAQRRV 9

RESULT 2  
 AAY05981  
 ID AAY05981 standard; Peptide; 9 AA.  
 XX  
 AC AAY05981;  
 XX  
 DT 16-AUG-1999 (first entry)  
 DE Human cancer antigen NY ESO-1/CAG-3 ORF2 cancer peptide.  
 XX  
 KW NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human; leukemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer; metastasis; melanoma; adenocarcinoma; thymoma; colon cancer; uterine cancer; breast cancer; prostate cancer; ovarian cancer; cervical cancer; bladder cancer; kidney cancer; pancreatic cancer; liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy; vaccine.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9918206-A2.  
 XX  
 PD 15-APR-1999.  
 XX  
 PF 21-SEP-1998; 98WO-US19609.  
 XX  
 PR 08-OCT-1997; 97US-0061428.  
 XX  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 PI Rosenberg SA, Wang RF;  
 XX  
 DR WPI; 1999-277270/23.  
 DR N-PSDB; AAX58602.  
 XX  
 XX Cancer antigen NY ESO1/CAG-3  
 PT  
 PS Claim 27; Page 65; 88pp; English.  
 XX  
 CC The present sequence represents a cancer peptide that corresponds to amino acid residues 19-27 of human ESO-1/CAG-3 (or CAG-3) ORF2 (see AAY05966), a new and potent tumour antigen capable of eliciting

an antigen specific immune response by T cells. Cancer peptides derived from CAG-3 ORF2, CAG-3 ORF1 (see AAY05965), portions of them and their variants (see AAY05967-87), are useful as cancer vaccines and that protect against cancer. The invention provides: vectors and host cells (also useful as vaccines); a method of diagnosis of cancer or precancer; a transgenic animal; antisense oligonucleotides that inhibit expression of the cancer peptide or tumour antigen; antibodies reacting with a CAG-3 cancer peptide, useful in diagnostic and detection assays; and methods for preventing or inhibiting cancer by administering a cancer peptide, with or without an HLA molecule. The cancer peptides form part of, or are derived from, cancers such as primary or metastatic melanoma, thymoma, lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine cancer, cervical cancer, bladder cancer, kidney cancer and adenocarcinomas such as breast, prostate, ovarian, pancreatic and thyroid cancers. Melanoma is treated by inducing cancer-specific T cells in vitro for subsequent return to a patient.

Query Match 78.0%; Score 32; DB 20; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAQRRV 9  
 |||||  
 Db 1 AAQRRV 7

RESULT 3  
 AAB09858  
 ID AAB09858 standard; peptide; 9 AA.  
 XX  
 AC AAB09858;  
 XX  
 DT 12-OCT-2000 (first entry)  
 DE Immunomodulatory peptide SEQ ID 27.  
 XX  
 KW Immunomodulatory peptide; AIDS; acquired immunodeficiency syndrome; infection; leukocytic disorder.  
 XX  
 OS Synthetic.  
 XX  
 PN US066622-A.  
 PD 23-MAY-2000.  
 XX  
 PF 28-OCT-1993; 93US-0144779.  
 XX  
 PR 28-OCT-1991; 91US-0783517.  
 PR 02-JAN-1992; 92US-0816205.  
 PR 28-OCT-1992; 92US-0967633.  
 XX  
 PA (CYTR-) CYTRAN INC.  
 XX  
 PI Green LR, Ivanov VT, Mikhalyova II, Vaskovsky BV, Mikhaltsov AN;  
 PI Khavinson VK, Morozov VG, Sinackevich NV;  
 XX  
 DR WPI; 2000-410680/35.  
 XX  
 PT Modulating immune system activity, useful for restoring the natural balance of the system and for treating diseases associated with suppressed or hyperactive immune systems, by administering a peptide with at least five amino acids -  
 PT  
 PS Disclosure; Column 37; 35pp; English.  
 XX  
 CC The present sequence is a peptide which can be used as an immunomodulator. This can be used to treat primary immunodeficiencies such as AIDS, DeGeorge's syndrome and severe combined immunodeficiency, secondary immunodeficiencies including anergy from tuberculosis, drug-induced leukopenia, non-HIV viral illnesses leukopenia, radiation

CC poisoning, toxin exposure and malnutrition, hyperactive immune states  
 CC including systemic lupus erythematosus, rheumatic fever, rheumatoid  
 CC arthritis and multiple sclerosis, bacterial, viral, fungal and parasitic  
 CC infections, and leukocytic disorders such as pre-leukaemias and leukemoid  
 CC reactions. Also, some anaemias may be treated, including acute  
 CC haemorrhagic anaemia, anaemias of chronic disease, megaloblastic  
 CC anaemias, iron deficiency anaemias and haemoglobinopathies.

SQ Sequence 9 AA;

Query Match 61.0%; Score 25; DB 21; Length 9;  
 Best Local Similarity 44.4%; Pred. No. 9.3e+05;  
 Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAAQERRV 9  
 |||:::|  
 Db 1 MLTAZZKAV 9

#### RESULT 4

ID AAE20695  
 AC AAE20695; standard; peptide; 9 AA.

XX

DT 01-JUL-2002 (first entry)

XX Immunomodulatory peptide #27 used to treat immune disorders.

DE Therapy; immunodeficient; hyperactive immune state; leprosy; infection;  
 KW spiriochetal infection; tuberculosis; human immunodeficiency virus; HIV;  
 KW aspergillosis; syphilis; Marek's disease; schistosomiasis; toxoplasmosis;  
 KW malaria; leishmaniasis; pneumocystis; leukocytic disorder; wound healing;  
 KW immune system; acquired immune deficiency syndrome; malnutrition; AIDS;  
 KW anaemia; tuberculosis; DeGeorge's syndrome; leukopaenia; toxin exposure;  
 KW radiation poisoning; multiple sclerosis; rheumatoid arthritis; vulvovaginitis;  
 KW cytostatic; dermatological; prophylaxis; systemic lupus erythematosus;  
 KW severe combined immunodeficiency; immunomodulatory peptide.

XX Unidentified.

OS US6346514-B1.

PN 12-FEB-2002.

PD 04-AUG-1999; 99US-0368449.

PF 28-OCT-1993; 93US-0144779.

PR 28-OCT-1991; 91US-0783517.

PR 02-JAN-1992; 92US-0816205.

PR 28-OCT-1992; 92US-0967633.

XX (CYTR-) CYTRAN INC.

XX Green LR, Sinackevich NV, Ivanov VT, Mikhalyova II, Vaskovsky BV;

PI Mikhailov AN, Khavinson VK, Morozov VG;

XX WPI; 2002-266497/31.

DR Treating bacterial, fungal and parasitic infections and immune

XX disorders in a host by administering an immunomodulatory peptide -

PS Disclosure; Column 37; 30pp; English.

XX The present invention relates to compositions and methods for the therapy

CC of immunodeficient, immunodepressed or hyperactive immune states. The

CC method of treating an infection in a host comprises administering an

CC immunomodulatory peptide to the host. The method is useful for treating

CC bacterial infections such as mycobacterial infections (e.g. tuberculosis,

CC leprosy), spirochetal infections (e.g. syphilis), viral infections such

CC and pneumocystis. They are also used to treat anaemias, atopic states,  
 CC leukocytic disorders (e.g. pre-leukaemias) and to enhance wound healing.  
 CC The method enhances or suppresses the immune system and immunological  
 CC enhancement is employed as a treatment for variety of primary diseases  
 CC (e.g., acquired immune deficiency syndrome (AIDS), DeGeorge's syndrome,  
 CC severe combined immunodeficiency), secondary immunodeficiencies (e.g.  
 CC anergy from tuberculosis and drug-induced leukopaenia), toxin exposure,  
 CC radiation poisoning, and malnutrition. Immunomodulation by the methods  
 CC of the invention is also useful for treating hyperactive immune states  
 CC such as systemic lupus erythematosus, rheumatoid arthritis and multiple  
 CC sclerosis. The immunomodulatory peptides of the invention can be  
 CC administered to a patient susceptible or at risk for infection, anaemia  
 CC or other disorders. They can be used for surgical prophylaxis to lessen  
 CC the risk of infectious complications and enhance the host's restorative  
 CC response to blood loss. The present sequence is an immunomodulatory  
 CC peptide of the invention.

SQ Sequence 9 AA;

Query Match 61.0%; Score 25; DB 23; Length 9;  
 Best Local Similarity 44.4%; Pred. No. 9.3e+05;  
 Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAAQERRV 9  
 |||:::|

Db 1 MLTAZZKAV 9

#### RESULT 5

ABP75370

ID ABP75370 standard; Peptide; 9 AA.

XX

AC ABP75370;

DT 20-FEB-2003 (first entry)

XX Chlamydia trachomatis peptide epitope #76.

DE Antibacterial; secreted protein; intracellular bacterium.

XX Chlamydia trachomatis.

OS WO200282091-A2.

PN 17-OCT-2002.

PD 09-APR-2002; 2002WO-DK00234.

PF 09-APR-2001; 2001DK-0000581.

PR 09-APR-2001; 2001US-282513P.

XX (SHAW/) SHAW A C.

XX (VAND/) VANDAH B B.

XX Shaw AC, Vandahl BB;

PI WPI; 2003-058585/05.

DR Identifying intracellular bacterial proteins by labeling proteins in

XX the presence of a eukaryotic protein synthesis inhibitor, performing

PT electrophoresis, autoradiography and comparing profiles to an

XX infected-cell lysate profile -

PS Claim 54; Page 174; 179pp; English.

XX The present invention relates to a method (M1) for identifying secreted

CC intracellular bacterial proteins (BP). M1 comprises: (a) selectively

CC visualising BP by pulse labelling in the presence of an inhibitor of

CC eukaryotic protein synthesis followed by 2D electrophoresis and

CC autoradiography; (b) comparing protein profiles (PF) of purified bacteria

CC to PF of total lysate (TL) of infected cells; and (c) identifying protein

CC spots present in differential images from gels loaded with TL. The

CC present sequence is one such bacterial peptide epitope which was

CC Identified by the method of the invention.

XX Sequence 9 AA;

Query Match 58.5%; Score 24; DB 24; Length 9;  
Best Local Similarity 75.0%; Pred. No. 9.3e+05;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 LAAQRRV 9

III I I I

DB 2 LAALNRV 9

# RESULT 6

ID AAB07254 standard; peptide; 7 AA.

AC AAB07254;

XX 12-OCT-2000 (first entry)

DT Immunomodulatory peptide SEQ ID 24.

DE Immunomodulatory peptide; AIDS; acquired immunodeficiency syndrome;  
infection; leukocytic disorder.

XX Synthetic.

XX US066622-A.

XX 23-MAY-2000.

XX 28-OCT-1993; 93US-0144779.

XX 28-OCT-1991; 91US-0783517.

PR 02-JAN-1992; 92US-0816205.

PR 28-OCT-1992; 92US-0967633.

XX (CYTR-) CYTRAN INC.

XX Green LR, Ivanov VT, Mikhailova II, Vaskovsky BV, Mikhailov AN;  
PI Khavinson VK, Morozov VG, Sinackevich NV;

XX WPI: 2000-410680/35.

XX Modulating immune system activity, useful for restoring the natural  
PT balance of the system and for treating diseases associated with  
PT suppressed or hyperactive immune systems, by administering a peptide  
PT with at least five amino acids -

XX Disclosure; Column 37; 35pp; English.

XX The present sequence is a peptide which can be used as an  
CC immunomodulator. This can be used to treat primary immunodeficiencies  
CC such as AIDS, DeGeorge's syndrome and severe combined immunodeficiency,  
CC secondary immunodeficiencies including energy from tuberculosis,  
CC drug-induced leukopenia, non-HIV viral illnesses leukopenia, radiation  
CC poisoning, toxin exposure and malnutrition, hyperactive immune states  
CC including systemic lupus erythematosus, rheumatic fever, rheumatoid  
CC arthritis and multiple sclerosis, bacterial, viral, fungal and parasitic  
CC infections, and leukocytic disorders such as pre-leukemias and leukemoid  
CC reactions. Also, some anaemias may be treated, including acute  
CC haemorrhagic anaemia, anaemias of chronic disease, megaloblastic  
CC anaemias, iron deficiency anaemias and haemoglobinopathies.

XX Sequence 7 AA;

Query Match 53.7%; Score 22; DB 21; Length 7;  
Best Local Similarity 42.9%; Pred. No. 9.3e+05;

Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MIAAQR 7

II I I I I

DB 1 MLTAZZK 7

# RESULT 7

AAE20692

ID AAE20692 standard; peptide; 7 AA.

XX AAE20692;

XX 01-JUL-2002 (first entry)

XX Immunomodulatory peptide #24 used to treat immune disorders.

XX Therapy; immunodeficient; hyperactive immune state; leprosy; infection;  
KW spirochetal infection; tuberculosis; human immunodeficiency virus; HIV;  
KW aspergillosis; syphilis; Marek's disease; schistosomiasis; toxoplasmosis;  
KW malaria; leishmaniasis; pneumocystis; leukocytic disorder; wound healing;  
KW immune system; acquired immune deficiency syndrome; malnutrition; AIDS;  
KW anaemia; tuberculosis; DeGeorge's syndrome; leukopenia; toxin exposure;  
KW radiation poisoning; multiple sclerosis; rheumatoid arthritis; vulvurary;  
KW cytostatic; dermatological; prophylaxis; systemic lupus erythematosus;  
KW severe combined immunodeficiency; immunomodulatory peptide.

XX Unidentified.

XX US6346514-B1.

XX 12-FEB-2002.

XX 04-AUG-1999; 99US-0368449.

XX 28-OCT-1993; 93US-0144779.

PR 28-OCT-1991; 91US-0783517.

PR 02-JAN-1992; 92US-0816205.

PR 28-OCT-1992; 92US-0967633.

XX (CYTR-) CYTRAN INC.

XX Green LR, Sinackevich NV, Ivanov VT, Mikhailova II, Vaskovsky BV;  
PI Mikhailov AN, Khavinson VK, Morozov VG;

XX WPI: 2002-266497/31.

XX Treating bacterial, fungal and parasitic infections and immune  
PT disorders in a host by administering an immunomodulatory peptide -

XX Disclosure; Column 37; 30pp; English.

XX The present invention relates to compositions and methods for the therapy  
CC of immunodeficient, immunodepressed or hyperactive immune states. The  
CC method of treating an infection in a host comprises administering an  
CC immunomodulatory peptide to the host. The method is useful for treating  
CC bacterial infections such as mycobacterial infections (e.g. tuberculosis,  
CC leprosy), spirochetal infections (e.g. syphilis), viral infections such  
CC as Marek's disease, HIV (human immunodeficiency virus) 1 and 2, fungal  
CC infections such as Candida albicans and aspergillosis and parasitic  
CC diseases such as malaria, schistosomiasis, toxoplasmosis, leishmaniasis  
CC and pneumocystis. They are also used to treat anaemias, atopic states,  
CC leukocytic disorders (e.g. pre-leukemias) and to enhance wound healing.  
CC The method enhances or suppresses the immune system and immunological  
CC enhancement is employed as a treatment for variety of primary diseases  
CC (e.g., acquired immune deficiency syndrome (AIDS), DeGeorge's syndrome,  
CC severe combined immunodeficiency), secondary immunodeficiencies (e.g.  
CC radiation poisoning, and malnutrition. Immunomodulation by the methods  
CC of the invention is also useful for treating hyperactive immune states  
CC such as systemic lupus erythematosus, rheumatoid arthritis and multiple  
CC sclerosis. The immunomodulatory peptides of the invention can be  
CC administered to a patient susceptible or at risk for infection, anaemia  
CC or other disorders. They can be used for surgical prophylaxis to lessen  
CC the risk of infectious complications and enhance the host's restorative  
CC response to blood loss. The present sequence is an immunomodulatory  
CC peptide of the invention.

XX SQ Sequence 7 AA;  
 Query Match 53.7%; Score 22; DB 23; Length 7;  
 Best Local Similarity 42.9%; Pred. No. 9.3e+05;  
 Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MLAAQER 7  
 |||:::  
 Db 1 MLTAZZK 7

RESULT 8  
 AAR36590  
 ID AAR36590 standard; peptide; 8 AA.  
 AC AAR36590;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 26-AUG-1993 (first entry)  
 XX  
 DE Immunodepressed state treatment peptide #1.  
 XX  
 KW Peptide; treatment; immunodepression; opportunistic infection; AIDS;  
 KW acquired immunodeficiency syndrome; pharmaceutical; salt; influenza;  
 KW restoration; stimulation; immune function; veterinary practice; anemia;  
 KW immunomodulatory; prophylaxis; hypotrophy; farming; animals; burns;  
 KW fur bearing; poultry; respiratory disease; wounds; open sores; rashes;  
 KW allergy; sun exposure; trauma; eczema; psoriasis; bone fractures;  
 KW lesions; gingival disease; gynecological; infection; infaralymphatic;  
 KW iron deficient.  
 XX  
 OS Synthetic.  
 OS WO9308816-A1.  
 PN  
 XX  
 PD 13-MAY-1993.  
 XX  
 PF 28-OCT-1992; 92WO-US09252.  
 XX  
 PR 28-OCT-1991; 91US-0783517.  
 XX  
 PA (CYTO-) CYTOVEN.  
 XX  
 PI Green LR;  
 XX  
 DR WPI; 1993-167379/20.  
 XX  
 PT Treatment of opportunistic infection in immuno:depressed patients -  
 PT by administering simple peptide(s)  
 XX  
 PS Disclosure; Page 3; 34pp; English.  
 XX  
 CC The sequences given in AAR36590-98 are simple peptides which can be  
 CC used in the treatment of immunodepressed states and of opportunistic  
 CC infections in immunodepressed states associated with acquired immuno-  
 CC deficiency syndrome. The peptides are all covered by the generic  
 CC sequence;  
 CC where R' = Thr-Ala, Thr-Pro, Ser-Ala, Ser-Pro, Ser-Ser, Met-Leu-Thr-  
 CC Ala, or Leu-Thr-Ala;  
 CC R" = -H, -Ala, -Ala-Ala or Ala-Val.  
 CC The peptides may be prepared by conventional methods of enzyme  
 CC synthesis and may be used as the active ingredients in a  
 CC pharmaceutical preparation as free peptides or in the form of a  
 CC water soluble salt, such as sodium, potassium, ammonium or zinc salt.  
 CC These peptides may be used for the restoration and stimulation of  
 CC immune functions. They may also be used in veterinary practice as an  
 CC immunomodulatory agent for the prophylaxis and treatment of hypo-  
 CC trophy in farming animals, fur bearing animals and poultry. The  
 CC opportunistic infections which may be treated using these peptides  
 CC include respiratory disease, influenza, AIDS, burns, wounds, other  
 CC open sores, rashes due to allergic reaction, sun exposure, local

CC trauma, eczema, psoriasis, etc. The peptides may be used to assist  
 CC healing in immunodepressed states such as healing bone fractures,  
 CC lesions, gingival diseases, gynecological infections, infaralymphatic  
 CC infections, etc. These peptides are particularly useful in the  
 CC treatment of anemias of all types such as those caused by loss of  
 CC blood, blood formation affection or blood destruction enhancement  
 CC including all types of iron deficient anemia.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 SQ Sequence 8 AA;  
 Query Match 53.7%; Score 22; DB 14; Length 8;  
 Best Local Similarity 42.9%; Pred. No. 9.3e+05;  
 Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MLAAQER 7  
 |||:::  
 Db 1 MLTAZZK 7

RESULT 9  
 AAB07255  
 ID AAB07255 standard; peptide; 8 AA.  
 XX  
 AC AAB07255;  
 XX  
 DT 12-OCT-2000 (first entry)  
 XX  
 DE Immunomodulatory peptide SEQ ID 25.  
 XX  
 KW Immunomodulatory peptide; AIDS; acquired immunodeficiency syndrome;  
 KW infection; leukocytic disorder.  
 XX  
 OS Synthetic.  
 OS US6066622-A.  
 PN  
 XX  
 PD 23-MAY-2000.  
 XX  
 PF 28-OCT-1993; 93US-0144779.  
 XX  
 PR 28-OCT-1991; 91US-0783517.  
 PR 02-JAN-1992; 92US-0816205.  
 PR 28-OCT-1992; 92US-0967633.  
 XX  
 PA (CYTR-) CYTRAN INC.  
 XX  
 PI Green LR, Ivanov VT, Mikhalyova II, Vaskovsky BV, Mikhaltsov AN;  
 PI Khavinson VK, Morozov VG, Sinackevich NV;  
 XX  
 DR WPI; 2000-410680/35.  
 XX  
 PT Modulating immune system activity, useful for restoring the natural  
 PT balance of the system and for treating diseases associated with  
 PT suppressed or hyperactive immune systems, by administering a peptide  
 PT with at least five amino acids -  
 XX  
 PS Disclosure; Column 37; 35pp; English.  
 XX  
 CC The present sequence is a peptide which can be used as an  
 CC immunomodulator. This can be used to treat primary immunodeficiencies  
 CC such as AIDS, DeGeorge's syndrome and severe combined immunodeficiency,  
 CC secondary immunodeficiencies including energy from tuberculosis,  
 CC drug-induced leukopenia, non-HIV viral illnesses leukopenia, radiation  
 CC poisoning, toxin exposure and malnutrition, hyperactive immune states  
 CC including systemic lupus erythematosus, rheumatic fever, rheumatoid  
 CC arthritis and multiple sclerosis, bacterial, viral, fungal and parasitic  
 CC infections, and leukocytic disorders such as pre-leukemias and leukemoid  
 CC reactions. Also, some anemias may be treated, including acute  
 CC haemorrhagic anaemia, anemias of chronic disease, megaloblastic  
 CC anemias, iron deficiency anaemias and haemoglobinopathies.

SQ Sequence 8 AA;

Query Match 53.7%; Score 22; DB 21; Length 8;  
 Best Local Similarity 42.9%; Pred. No. 9.3e+05;  
 Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAAQER 7  
 |||:::  
 Db 1 MLTAZZK 7

## RESULT 10

AAE20693  
 ID AAE20693 standard; peptide; 8 AA.

AC AAE20693;

DT 01-JUL-2002 (first entry)

DE Immunomodulatory peptide #25 used to treat immune disorders.

XX Therapy: immunodeficient; hyperactive immune state; leprosy; infection;  
 KW spirochetal infection; tuberculosis; human immunodeficiency virus; HIV;  
 KW aspergillosis; syphilis; Marek's disease; schistosomiasis; toxoplasmosis;  
 KW malaria; leishmaniasis; pneumocystis; leukocytic disorder; wound healing;  
 KW immune system; acquired immune deficiency syndrome; malnutrition; AIDS;  
 KW anaemia; tuberculosis; DeGeorge's syndrome; leukopaenia; toxin exposure;  
 KW radiation poisoning; multiple sclerosis; rheumatoid arthritis; vulnery;  
 KW cytostatic; dermatological; prophylaxis; systemic lupus erythematosus;  
 KW severe combined immunodeficiency; immunomodulatory peptide.

XX Unidentified.

XX US6346514-B1.

XX 12-FEB-2002.

XX 04-AUG-1999; 99US-0368449.

XX 28-OCT-1993; 93US-0144779.

XX 28-OCT-1991; 91US-0783517.

XX 02-JAN-1992; 92US-0816205.

XX 28-OCT-1992; 92US-0967633.

XX (CYTR-) CVTRAN INC.

XX Green LR, Sinackevich NV, Ivanov VT, Mikhalyova II, Vaskovsky BV;

XX Mikhailtsov AN, Khavinson VK, Morozov VG;

XX WPI; 2002-266497/31.

XX Treating bacterial, fungal and parasitic infections and immune

XX disorders in a host by administering an immunomodulatory peptide -

XX Disclosure; Column 37; 30pp; English.

XX The present invention relates to compositions and methods for the therapy  
 CC of immunodeficient, immunodepressed or hyperactive immune states. The  
 CC method of treating an infection in a host comprises administering an  
 CC immunomodulatory peptide to the host. The method is useful for treating  
 CC bacterial infections such as mycobacterial infections (e.g. tuberculosis,  
 CC leprosy), spirochetal infections (e.g. syphilis), viral infections such  
 CC as Marek's disease, HIV (human immunodeficiency virus) 1 and 2, fungal  
 CC infections such as Candida albicans and aspergillosis and parasitic  
 CC diseases such as malaria, schistosomiasis, toxoplasmosis, leishmaniasis  
 CC and pneumocystis. They are also used to treat anaemias, atopic states,  
 CC leukocytic disorders (e.g. pre-leukaemias) and to enhance wound healing.  
 CC The method enhances or suppresses the immune system and immunological  
 CC enhancement is employed as a treatment for variety of primary diseases  
 CC (e.g., acquired immune deficiency syndrome (AIDS), DeGeorge's syndrome,  
 CC severe combined immunodeficiency), secondary immunodeficiencies (e.g.  
 CC energy from tuberculosis and drug-induced leukopaenia), toxin exposure,  
 CC radiation poisoning, and malnutrition. Immunomodulation by the methods

CC of the invention is also useful for treating hyperactive immune states  
 CC such as systemic lupus erythematosus, rheumatoid arthritis and multiple  
 CC sclerosis. The immunomodulatory peptides of the invention can be  
 CC administered to a patient susceptible or at risk for infection, anaemia  
 CC or other disorders. They can be used for surgical prophylaxis to lessen  
 CC the risk of infectious complications and enhance the host's restorative  
 CC response to blood loss. The present sequence is an immunomodulatory  
 CC peptide of the invention.

XX SQ Sequence 8 AA;

Query Match 53.7%; Score 22; DB 23; Length 8;  
 Best Local Similarity 42.9%; Pred. No. 9.3e+05;  
 Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAAQER 7  
 |||:::  
 Db 1 MLTAZZK 7

## RESULT 11

AAE09857  
 ID AAE09857 standard; peptide; 9 AA.

AC AAE09857;

XX 12-OCT-2000 (first entry)

XX Immunomodulatory peptide SEQ ID 26.

XX Immunomodulatory peptide; AIDS; acquired immunodeficiency syndrome;  
 KW infection; leukocytic disorder.

XX Synthetic.

XX US6066622-A.

XX 23-MAY-2000.

XX 28-OCT-1993; 93US-0144779.

XX 28-OCT-1991; 91US-0783517.

XX 02-JAN-1992; 92US-0816205.

XX 28-OCT-1992; 92US-0967633.

XX (CYTR-) CYTRAN INC.

XX Green LR, Ivanov VT, Mikhalyova II, Vaskovsky BV, Mikhailtsov AN;

XX Khavinson VK, Morozov VG, Sinackevich NV;

XX WPI; 2000-410680/35.

XX Modulating immune system activity, useful for restoring the natural  
 PT balance of the system and for treating diseases associated with  
 PT suppressed or hyperactive immune systems, by administering a peptide  
 PT with at least five amino acids -

XX Disclosure; Column 37; 35pp; English.

XX The present sequence is a peptide which can be used as an  
 CC immunomodulator. This can be used to treat primary immunodeficiencies  
 CC such as AIDS, DeGeorge's syndrome and severe combined immunodeficiency,  
 CC secondary immunodeficiencies including energy from tuberculosis,  
 CC drug-induced leukopaenia, non-HIV viral illnesses leukopaenia, radiation  
 CC poisoning, toxin exposure and malnutrition, hyperactive immune states  
 CC including systemic lupus erythematosus, rheumatic fever, rheumatoid  
 CC arthritis and multiple sclerosis, bacterial, viral, fungal and parasitic  
 CC infections, and leukocytic disorders such as pre-leukaemias and leukemoid  
 CC reactions. Also, some anaemias may be treated, including acute  
 CC haemorrhagic anaemia, anaemias of chronic disease, megaloblastic  
 CC anaemias, iron deficiency anaemias and haemoglobinopathies.

XX SQ Sequence 9 AA;

Query Match 53.7%; Score 22; DB 21; Length 9;  
Best Local Similarity 42.9%; Pred. No. 9.3e+05;  
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MIAQER 7  
|||:::  
DB 1 MLTAZZK 7

RESULT 12  
AAB69558  
ID AAB69558 standard; Peptide; 9 AA.

XX AC AAB69558;  
DT 27-APR-2001 (first entry)

XX DE Human Repro-EN-1.0 hydrophilic peptide #1.

XX KW Human; Repro-EN-1.0; IB1; cytostatic; breast cancer; uterine cancer;  
KW prostate cancer; epitope.

XX OS Homo sapiens.

XX WO200107616-A1.

XX PD 01-FEB-2001.

XX PF 10-MAR-2000; 2000WO-US06742.

XX PR 22-JUL-1999; 99US-0359084.

XX PR 30-JUL-1999; 99WO-US17284.

XX PR 23-NOV-1999; 99US-0447399.

XX PA (DIAG-) DIAGNOSTIC PROD CORP.

XX PI El Shaml AS, Menon SN, French CK;

XX WIPI; 2001-182795/18.

XX PT New autoantigens Repro-EN-1.0 and IB 1 polypeptides and polynucleotides  
are useful for diagnosing endometriosis or as a marker for pathologic  
conditions such as breast, uterine or prostate cancer -

XX PS Disclosure; Page 48; 119pp; English.

XX CC The present sequence is given in a specification relating to recombinant  
polynucleotides comprising nucleotide sequences encoding a polypeptide  
epitope of at least 5 amino acids of Repro-EN-1.0 or of IB 1 having a  
fully defined sequence of 860 and 937 amino acids. The epitope  
specifically binds to antibodies from subjects diagnosed with  
endometriosis. The Repro-EN-1.0 and IB 1 proteins are useful for  
diagnosing endometriosis by detecting antibodies in immunoassays, and  
are used as markers for pathologic conditions e.g. breast, uterine or  
prostate cancer. Methods for detecting Repro-EN-1.0 or IB 1  
polynucleotides or polypeptides are useful in the diagnosis of these  
cancers, monitoring their progress or treatment, and determining patient  
prognosis. Fragments of polynucleotides encoding Repro-EN-1.0 and IB 1  
may be used as probes for detecting mRNA from cell types suspected to  
be cancerous, and as primers for amplifying sequences. Repro-EN-1.0 and  
IB 1 polypeptides and immunogenic fragments may be used as positive  
controls in diagnostic assays to detect antibodies that specifically  
bind the proteins from patient serum samples. The polypeptides are  
useful as immunogens for eliciting antibody production against epitopes  
of the protein, and as controls in diagnostic methods.

XX SQ Sequence 9 AA;

Query Match 53.7%; Score 22; DB 22; Length 9;  
Best Local Similarity 66.7%; Pred. No. 9.3e+05;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAQERR 8  
:|::|:  
DB 4 SAEERR 9

RESULT 13  
AAE20694  
ID AAE20694 standard; peptide; 9 AA.

XX AC AAE20694;

XX DT 01-JUL-2002 (first entry)

XX DE Immunomodulatory peptide #26 used to treat immune disorders.

XX KW Therapy; immunodeficient; hyperactive immune state; leprosy; infection;  
KW spirochetal infection; tuberculosis; human immunodeficiency virus; HIV;  
KW aspergillosis; syphilis; Marek's disease; schistosomiasis; toxoplasmosis;  
KW malaria; leishmaniasis; pneumocystis; leukocytic disorder; wound healing;  
KW immune system; acquired immune deficiency syndrome; malnutrition; AIDS;  
KW anaemia; tuberculosis; DeGeorge's syndrome; leukopaenia; toxin exposure;  
KW radiation poisoning; multiple sclerosis; rheumatoid arthritis; vulnery;  
KW cytostatic; dermatological; prophylaxis; systemic lupus erythematosus;  
KW severe combined immunodeficiency; immunomodulatory peptide.

XX OS Unidentified.

XX US6346514-B1.

XX PD 12-FEB-2002.

XX PF 04-AUG-1999; 99US-0368449.

XX PR 28-OCT-1993; 93US-0144779.

XX PR 28-OCT-1991; 91US-0783517.

XX PR 02-JAN-1992; 92US-0816205.

XX PR 28-OCT-1992; 92US-0967633.

XX (CYTR-) CYTRAN INC.

XX PI Green LR, Sinackevich NV, Ivanov VT, Mikhalyova II, Vaskovsky BV;

XX PI Mikhailtsov AN, Khavinson VK, Morozov VG;

XX WIPI; 2002-266497/31.

XX PT Treating bacterial, fungal and parasitic infections and immune  
disorders in a host by administering an immunomodulatory peptide -

XX PS Disclosure; Column 37; 30pp; English.

XX CC The present invention relates to compositions and methods for the therapy  
of immunodeficient, immunodepressed or hyperactive immune states. The  
method of treating an infection in a host comprises administering an  
immunomodulatory peptide to the host. The method is useful for treating  
bacterial infections such as mycobacterial infections (e.g. tuberculosis,  
leprosy), spirochetal infections (e.g. syphilis), viral infections such  
as Marek's disease, HIV (human immunodeficiency virus) 1 and 2, fungal  
infections such as Candida albicans and aspergillosis and parasitic  
diseases such as malaria, schistosomiasis, toxoplasmosis, leishmaniasis  
and pneumocystis. They are also used to treat anaemias, atopic states,  
leukocytic disorders (e.g. pre-leukaemias) and to enhance wound healing.  
The method enhances or suppresses the immune system and immunological  
enhancement is employed as a treatment for variety of primary diseases  
(e.g., acquired immune deficiency syndrome (AIDS), DeGeorge's syndrome,  
severe combined immunodeficiency), secondary immunodeficiencies (e.g.  
anergy from tuberculosis and drug-induced leukopaenia), toxin exposure,  
radiation poisoning, and malnutrition. Immunomodulation by the methods  
of the invention is also useful for treating hyperactive immune states  
such as systemic lupus erythematosus, rheumatoid arthritis and multiple  
sclerosis. The immunomodulatory peptides of the invention can be  
administered to a patient susceptible or at risk for infection, anaemia  
or other disorders. They can be used for surgical prophylaxis to lessen  
the risk of infectious complications and enhance the host's restorative

CC response to blood loss. The present sequence is an immunomodulatory  
CC peptide of the invention.

XX  
SQ Sequence 9 AA;

Query Match 53.7%; Score 22; DB 23; Length 9;  
Best Local Similarity 42.9%; Pred. No. 9.3e+05;  
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAQER 7  
|||:  
Db 1 MLTAZK 7

## RESULT 14

AAV23550  
ID AAV23550 standard; Protein; 6 AA.

XX AC AAV23550;

DT 02-SEP-1999 (first entry)

XX T-cell receptor V beta 7 CDR3 region sequence.

XX Vaccine: T cell receptor; TCR; T cell; V beta 6.2/3; V beta 6/5;  
KW V beta 6.7; V beta 2; V beta 5/1; V beta 7; V beta 13; V beta 6;  
KW multiple sclerosis.

XX Homo sapiens.

XX W09927957-A1.

XX 10-JUN-1999.

PF 03-DEC-1997; 97WO-US23147.

PR 03-DEC-1997; 97WO-US23147.

XX (IMMU-) IMMUNE RESPONSE CORP.

PA (KIMM-) KIMMEL CANCER CENT SIDNEY.

PI Brostoff SW, Carlo DJ, Gold DP, Smith LR, Wilson DB;

DR WPI; 1999-404801/34.

XX T0 cell receptor peptide-derived vaccines

PS Disclosure; Fig 3; 104pp; English.

XX The specification describes vaccines which comprise immunologically  
CC effective amounts of T cell receptor (TCR) peptides. The TCRs are  
CC present on the surface of T cells. The TCRs are chosen from V beta  
CC 6.2/3, V beta 6/5, V beta 6.7, V beta 2, V beta 5/1, V beta 7 or V beta  
CC 13. The V beta TCR peptide-based vaccines are useful for prevention or  
CC treatment of multiple sclerosis (MS). The presence of V beta 6.7 appears  
CC to be particularly associated with multiple sclerosis and can be used  
CC to determine an individual's susceptibility to multiple sclerosis.  
CC Vaccinating, rather than passively administering heterologous  
CC antibodies, allows the host's own immune system to mobilize and suppress  
CC auto aggressive T cells. Therefore, the suppression is persistent and  
CC may involve any and all immunological mechanisms in effecting that  
CC suppression. Such a multi-faceted response is more effective than  
CC the uni-dimensional suppression achieved by passive administration of  
CC monoclonal antibodies or extant-derived regulatory T cell clones.  
CC AAV23550-59 represent CDR3 regions of monoclonal and oligoclonal TCR  
CC V beta 7 clones from MS patients, grown in IL-2 and IL-4.

XX Sequence 6 AA;

Query Match 51.2%; Score 21; DB 20; Length 6;  
Best Local Similarity 80.0%; Pred. No. 9.3e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 QERRV 9  
|:|:  
Db 1 QDRRV 5

## RESULT 15

AAB80661  
ID AAB80661 standard; peptide; 7 AA.

XX AC AAB80661;

XX 02-MAY-2001 (first entry)

XX Human glandular kallikrein 2 cleavage site peptide #22.

XX Cleavage; kallikrein 2; hK2; prodrug.

XX Synthetic.

XX W0200109165-A2.

XX 08-FEB-2001.

XX 28-JUL-2000; 2000WO-US40496.

XX 29-JUL-1999; 99US-0146316.

XX (UYJO ) UNIV JOHNS HOPKINS.

XX Denmeade SR, Isaacs JT, Lilja H, Christensen SB;

XX WPI; 2001-191450/19.

XX New peptides containing cleavage sites specifically cleaved by human  
XX kallikrein 2, useful for producing prodrugs which treat hK2-producing  
XX cell proliferative disorders without exhibiting non-specific toxicity

XX Disclosure; Page 8; 38pp; English.

XX The present invention relates to a peptide comprising an amino acid  
XX sequence having a cleavage site specific for an enzyme having a  
XX proteolytic activity of human kallikrein 2 (hK2), and which is up to  
XX 20 amino acids in length. The invention is useful for producing a  
XX prodrug which involves linking a drug which contains a primary amine  
XX to the peptide, in which the linking of the peptide to the drug  
XX inhibits the therapeutic activity of the drug.

XX Sequence 7 AA;

Query Match 51.2%; Score 21; DB 22; Length 7;  
Best Local Similarity 66.7%; Pred. No. 9.3e+05;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 AQERRV 9  
|||:  
Db 2 AQKRL 7

Search completed: October 7, 2003, 14:45:32  
Job time : 57.5 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run On: October 7, 2003, 14:42:06 ; Search time 24.5 Seconds  
(without alignments)  
35.327 Million cell updates/sec

Title: US-09-807-512-26  
Perfect score: 41  
Sequence: 1 MLAGERRV 9  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues  
Total number of hits satisfying chosen parameters: 789

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_76:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17	41.5	8	4 I55411	hypothetical histo
2	15	36.6	6	2 A49792	acylaminoacyl-pept
3	15	36.6	8	2 A54823	olfactory receptor
4	15	36.6	8	2 B54823	olfactory receptor
5	15	36.6	8	2 I64832	Ca2+-transporting
6	14	34.1	7	2 PH0932	T-cell receptor be
7	14	34.1	9	2 A44873	caldesmon - rabbit
8	13	31.7	6	2 A19780	transferrin - bovi
9	13	31.7	8	2 S70727	ip3F protein - Shi
10	13	31.7	8	2 I48935	apolipoprotein A-I
11	13	31.7	9	2 S02384	probable membrane
12	13	31.7	9	2 S66608	quinoline 2-oxidor
13	12	29.3	5	2 S55237	zinc-binding prote
14	12	29.3	5	2 PT0672	T-cell receptor be
15	12	29.3	6	2 PC4127	hypothetical 6 pro
16	12	29.3	6	2 PT0510	T-cell receptor be
17	12	29.3	7	2 PS0254	18K protein 5507 -
18	12	29.3	8	2 PT0323	Ig heavy chain CRD
19	12	29.3	9	2 A43848	cell surface adhes
20	12	29.3	9	2 D57444	neuropeptide Grb-A
21	12	29.3	9	2 A45199	L-lysophosphorin - Ja
22	12	29.3	9	2 C60070	gastrin - domestic
23	12	29.3	9	2 A33527	fructose-2,6-bisph
24	12	29.3	9	4 S15594	orf 1 rara 5'-regi
25	11	26.8	5	2 I50385	myosin light chain
26	11	26.8	6	2 A35890	RNA-directed DNA p
27	11	26.8	6	2 I48126	alpha-tubulin - Ch
28	11	26.8	7	2 PT0676	T-cell receptor be
29	11	26.8	8	2 A32523	peptidyl-dipeptida

30	11	26.8	8	2 T10077	hypothetical prote
31	11	26.8	8	2 PC4131	hypothetical prote
32	11	26.8	8	2 PQ0726	unidentified 4.5/4
33	11	26.8	8	2 PT0691	T-cell receptor be
34	11	26.8	8	2 I57532	gene TnIsIow prote
35	11	26.8	9	2 SI3889	phosphoenolpyruvat
36	11	26.8	9	2 D44787	calliFMRFamide 13
37	11	26.8	9	2 S78762	ribosomal protein
38	11	26.8	9	2 PC7076	spectrin alpha cha
39	10	24.4	3	3 A33802	thyrotropin-releas
40	10	24.4	4	2 I40505	hypothetical prote
41	10	24.4	5	2 A26830	mitosis inhibiting
42	10	24.4	5	2 D44823	synaptosomal-assoc
43	10	24.4	5	2 PT0610	T-cell receptor be
44	10	24.4	5	2 PT0513	T-cell receptor be
45	10	24.4	5	2 PT0659	T-cell receptor be

ALIGNMENTS

RESULT 1

I55411  
hypothetical histone H2A.X (mistranslated) - human (fragment)  
C:Species: Homo sapiens (man)  
C>Date: 28-Jan-2000 #sequence\_revision 28-Jan-2000 #text\_change 28-Jan-2000  
C:Accession: I55411  
R:Ivanova, V.S.; Hatch, C.L.; Bonner, W.M.  
J. Biol. Chem. 269, 24189-24194, 1994  
A:Title: Characterization of the human histone H2A.X gene. Comparison of its promoter  
A:Reference number: I55411; MUID:95014156; PMID:7929075  
A:Accession: I55411  
A>Status: translation not shown; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-8 <IVA>  
A:Cross-references: GB:S73863; NID:Q765295; PIDN:AAD14141.1; PID:94261841  
A>Note: This is a hypothetical translation of a sequence from the promoter region pro  
C:Genetics:  
A:Gene: H2A.X

Query Match 41.5%; Score 17; DB 4; Length 8;  
Best Local Similarity 66.7%; Pred. No. 2.8e+05;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAQQR 7  
DB 3 LAARR 8

RESULT 2

A49792  
acylaminoacyl-peptidase (EC 3.4.19.1) - rabbit (fragment)  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C>Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 07-May-1999  
C:Accession: A49792  
R:Krishna, R.G.; Chin, C.C.Q.; Wold, F.  
Anal. Biochem. 199, 45-50, 1991  
A:Title: N-terminal sequence analysis of N(alpha)-acetylated proteins after unblockin

A:Reference number: A49792; MUID:92222120; PMID:1807161  
A:Accession: A49792  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-6 <KRI>

C:Keywords: acetylated amino end; hydrolase; omega peptidase  
F:I/Modified site: acetylated amino end (Met) #status experimental

Query Match 36.6%; Score 15; DB 2; Length 6;  
Best Local Similarity 75.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 ERRV 9  
DB 2 ERQV 5

RESULT 3  
A54823  
olfactory receptor I7 - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C>Date: 28-Apr-1995 #sequence\_revision 28-Apr-1995 #text\_change 17-Mar-1999  
C:Accession: A54823  
R:Chess, A.; Simon, I.; Cedar, H.; Axel, R.  
Cell 78, 823-834, 1994  
A:Title: Allelic inactivation regulates olfactory receptor gene expression.  
A:Reference number: A54823; MUID:94373818; PMID:8087849  
A:Accession: A54823  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-8 <CHE>

Query Match 36.6%; Score 15; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ERR 8  
Db 2 ERR 4

RESULT 4  
B54823  
olfactory receptor I7 - western wild mouse (fragment)  
C:Species: Mus spretus (western wild mouse)  
C>Date: 28-Apr-1995 #sequence\_revision 28-Apr-1995 #text\_change 17-Mar-1999  
C:Accession: B54823  
R:Chess, A.; Simon, I.; Cedar, H.; Axel, R.  
Cell 78, 823-834, 1994  
A:Title: Allelic inactivation regulates olfactory receptor gene expression.  
A:Reference number: A54823; MUID:94373818; PMID:8087849  
A:Accession: B54823  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-8 <CHE>

Query Match 36.6%; Score 15; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ERR 8  
Db 2 ERR 4

RESULT 5  
I64832  
Ca2+-transporting ATPase (EC 3.6.3.8) - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 19-Apr-2002  
C:Accession: I64832  
R:Wu, K.  
Am. J. Physiol. 264, 333-341, 1993  
A:Title: Molecular cloning and quantification of mRNA and protein encoding Ca2+-ATPase 1  
A:Reference number: I51892  
A:Accession: I64832  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 1-8 <RES>  
A:Cross-references: GB:M99223; NID:g203644; PIDN:AAA40992.1; PID:g203646  
C:Genetics:  
A:Gene: SERCALB  
C:Keywords: Hydrolase

Query Match 36.6%; Score 15; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ERR 8  
Db 5 ERR 7

RESULT 6  
PH0932  
T-cell receptor beta chain V-D-J region (clone 3) - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 09-Oct-1992 #sequence\_revision 09-Oct-1992 #text\_change 30-May-1997  
C:Accession: PH0932  
R:Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.  
J. Exp. Med. 174, 1467-1476, 1991  
A:Title: Analysis of T cell receptor beta chains in Lewis rats with experimental alle  
A:Reference number: PH0891; MUID:92078857; PMID:1836012  
A:Accession: PH0932  
A:Molecule type: mRNA  
A:Residues: 1-7 <GOI>  
A:Experimental source: complete Freund's adjuvant-immunized lymph node  
C:Keywords: T-cell receptor

Query Match 34.1%; Score 14; DB 2; Length 7;  
Best Local Similarity 60.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 AAOER 7  
Db 2 ASPER 6

RESULT 7  
A44873  
caldesmon - rabbit (fragment)  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C>Date: 31-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 31-Oct-1997  
C:Accession: A44873  
R:Ikebe, M.; Hornick, T.  
Arch. Biochem. Biophys. 288, 538-542, 1991  
A:Title: Determination of the phosphorylation sites of smooth muscle caldesmon by pro  
A:Reference number: A44873; MUID:91378498; PMID:1898046  
A:Accession: A44873  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-9 <IKE>  
A:Experimental source: skeletal myosin  
A>Note: sequence extracted from NCBI backbone (NCBIP:63199)  
C:Superfamily: caldesmon

Query Match 34.1%; Score 14; DB 2; Length 9;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LAAQER 7  
Db 4 LXIEER 9

RESULT 8  
A19780  
transferrin - bovine (fragment)  
C:Species: Bos primigenius taurus (cattle)  
C>Date: 25-Oct-1987 #sequence\_revision 25-Oct-1987 #text\_change 18-Jun-1993  
C:Accession: A19780  
R:Brook, J.H.; Esparza, I.; Oliver, R.A.; Spooner, R.L.  
Biochem. Genet. 18, 851-860, 1980  
A:Title: Electrophoretic mobility of N- and C-terminal monoferric fragments of bovine  
A:Reference number: A19780; MUID:81183891; PMID:7225082  
A:Accession: A19780  
A:Molecule type: protein  
A:Residues: 1-6 <BRO>

Query Match 31.7%; Score 13; DB 2; Length 6;  
Best Local Similarity 75.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 ERRV 9  
|||  
Db 3 ERTV 6

## RESULT 9

S70727  
IpgF protein - Shigella flexneri (fragment)  
C:Species: Shigella flexneri  
C:Date: 15-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 08-Oct-1999  
C:Accession: S70727  
R:Aliaoui, A.; Sansonetti, P.J.; Menard, R.; Barzu, S.; Mounier, J.; Phalipon, A.; Parsot, M.; Microbiol. 17, 461-470, 1995  
A:Title: MxIG, a membrane protein required for secretion of Shigella spp. Ipa invasins.  
A:Reference number: S70727; MUID:96100445; PMID:8559065  
A:Accession: S70727  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-8 <ALL>  
A:Cross-references: EMBL:Z48957; NID:g929880; PIDN:CAAB8821.1; PID:g929881  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1995  
C:Genetics:  
A:Gene: ipgF

Query Match 31.7%; Score 13; DB 2; Length 8;  
Best Local Similarity 66.7%; Pred. No. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 RRV 9  
|||  
Db 6 RRI 8

## RESULT 10

I48935  
apolipoprotein A-II - western wild mouse (fragment)  
C:Species: Mus spretus (western wild mouse)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 05-Nov-1999  
C:Accession: I48935  
R:Ko, M.S.; Wang, X.; Horton, J.H.; Hagen, M.D.; Takahashi, N.; Maezaki, Y.; Nadeau, J.H. Mamm. Genome 5, 349-355, 1994  
A:Title: Genetic mapping of 40 cDNA clones on the mouse genome by PCR.  
A:Reference number: I48935; MUID:94319082; PMID:8043949  
A:Accession: I48935  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-8 <RES>  
A:Cross-references: EMBL:U05692; NID:g497011; PIDN:AAB60463.1; PID:g642827

Query Match 31.7%; Score 13; DB 2; Length 8;  
Best Local Similarity 75.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAAQ 5  
|||  
Db 5 LAAK 8

## RESULT 11

S02384  
probable membrane antigen CL4 - human herpesvirus 4 (fragment)  
C:Species: human herpesvirus 4, Epstein-Barr virus  
C:Date: 01-Dec-1989 #sequence\_revision 01-Dec-1989 #text\_change 08-Oct-1999  
C:Accession: S02384  
R:Walls, D.; Gannon, F.; EMBO J. 7, 1191-1196, 1988  
A:Title: The expression of novel antigens from the Epstein-Barr virus large internal repeat region.  
A:Reference number: S02381; MUID:88296424; PMID:2841116  
A:Accession: S02384  
A:Molecule type: DNA  
A:Residues: 1-9 <NAL>

A:Cross-references: EMBL:X07817; NID:g59173; PIDN:CAA30676.1; PID:g930015  
C:Keywords: membrane protein; surface antigen

Query Match 31.7%; Score 13; DB 2; Length 9;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LAAQER 7  
|: ||  
Db 3 LSLAER 8

## RESULT 12

S66608  
quinoline 2-oxidoeductase gamma chain - Comamonas testosteroni (fragment)  
C:Species: Comamonas testosteroni  
C:Date: 15-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 17-Mar-1999  
C:Accession: S66608  
R:Schach, S.; Tshisuaka, B.; Fetzner, S.; Lingens, F. Eur. J. Biochem. 232, 536-544, 1995  
A:Title: Quinoline 2-oxidoeductase and 2-oxo-1,2-dihydroquinoline 5,6-dioxygenase from *Comamonas testosteroni*.  
A:Reference number: S66608; MUID:96035889; PMID:7556204  
A:Accession: S66608  
A:Molecule type: protein  
A:Residues: 1-9 <SCH>  
A:Experimental source: strain 63

Query Match 31.7%; Score 13; DB 2; Length 9;  
Best Local Similarity 33.3%; Pred. No. 2.8e+05;  
Matches 2; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAAQE 6  
|: |::  
Db 1 MIQAEE 6

## RESULT 13

S55237  
zinc-binding protein ZBP14 - maize (fragment)  
C:Species: Zea mays (maize)  
C:Date: 27-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 07-May-1999  
C:Accession: S55237  
R:Robinson, K.; Jones, D.; Howell, S.; Soneji, Y.; Martin, S.; Aitken, A. Biochem. J. 307, 267-272, 1995  
A:Title: Expression and characterization of maize ZBP14, a member of a new family of zinc-binding proteins.  
A:Reference number: S55237; MUID:95234046; PMID:7717986  
A:Accession: S55237  
A:Molecule type: protein  
A:Residues: 1-5 <ROB>

Query Match 29.3%; Score 12; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAA 4  
|||  
Db 1 LAA 3

## RESULT 14

PT0672  
T-cell receptor beta chain V-D-J region (121-1B) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
C:Accession: PT0646; PT0672  
R:Feeney, A.J.; J. Exp. Med. 174, 115-124, 1991  
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
A:Reference number: PT0509; MUID:91277601; PMID:1711558  
A:Accession: PT0646  
A>Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-5 <FEE>

A: Experimental source: day 4 postnatal thymus, strain BALB/c, clone 121-1B  
A: Accession: PT0672  
A: Status: translation not shown  
A: Molecule type: DNA  
A: Residues: 1-5 <FE2>  
A: Experimental source: day 18 fetal thymus, strain BALB/c, clone 140-1BG  
C: Keywords: T-cell receptor

Query Match 29.3%; Score 12; DB 2; Length 5;  
Best Local Similarity 40.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 AAQER 7  
|: :|  
Db 1 ASSDR 5

## RESULT 15

PC4127  
hypothetical 6 protein - Streptomyces clavuligerus (fragment)  
C: Species: Streptomyces clavuligerus  
C: Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 16-Jul-1999  
C: Accession: PC4127  
R: Rodriguez-Garcia, A.; Martin, J.F.; Liras, P.  
Gene 167, 9-15, 1995  
A: Title: The argG gene of Streptomyces clavuligerus has low homology to unstable argG fr  
A: Reference number: J04548; MUID:96144242; PMID:8566818  
A: Accession: PC4127  
A: Status: preliminary  
A: Molecule type: DNA  
A: Residues: 1-6 <ROD>  
A: Cross-references: EMBL:Z49111

Query Match 29.3%; Score 12; DB 2; Length 6;  
Best Local Similarity 60.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 AQERR 8  
|||  
Db 1 AQNDR 5

Search completed: October 7, 2003, 14:49:05  
Job time : 25.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:32:11 ; Search time 13 Seconds  
(without alignments)  
32.557 Million cell updates/sec

Title: US-09-807-512-26

Perfect score: 41

Sequence: 1 MAAQERRV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 251

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	20	48.8	8	UH09_RAT	P56575 rattus norv
2	16	39.0	9	ULAK_MOUSE	P99031 mus musculu
3	15	36.6	6	ACPH_RABIT	P25154 oryctolagus
4	14	34.1	8	B44K_PORGI	P81886 porphyromon
5	14	34.1	8	CLP_THICU	P80488 thiobacillu
6	14	34.1	9	RS10_SERMA	O68936 serratia ma
7	13	31.7	8	UPAL_HUMAN	P30087 homo sapien
8	12	29.3	7	UH11_RAT	P56576 rattus norv
9	12	29.3	8	NS3_MYCTU	P81152 mycobacteri
10	11	26.8	9	FARD_CALVO	P41868 calliphora
11	11	26.8	9	ULAH_HUMAN	P31934 homo sapien
12	10	24.4	7	CIA_ENTFA	P11932 enterococcu
13	10	24.4	8	NPB_BOVIN	P15507 bos taurus
14	10	24.4	8	UC26_MAIZE	P80632 zea mays (m
15	10	24.4	9	FIBB_PAPPA	P19344 papio anubi
16	10	24.4	9	FIBB_PAPHA	P19343 papio hamad
17	10	24.4	9	FIBB_THEGE	P19342 theropithec
18	10	24.4	9	NEUX_HUMAN	P04277 homo sapien
19	9	22.0	7	CARP_MYTED	P10420 mytilus edu
20	9	22.0	7	IGAO_DACDE	P36294 dactylium d
21	9	22.0	7	LANC_CARUI	P36960 carnobacter
22	9	22.0	8	ANG2_BOTJA	Q10582 bothrops ja
23	9	22.0	8	CAD1_ENTFA	P13268 enterococcu
24	9	22.0	8	LPMS_STAEP	P23211 staphylococ
25	9	22.0	9	BS43_SERPL	P83375 serratia pl
26	9	22.0	9	FAR3_PENMO	P83318 penaeus mon
27	9	22.0	9	HUTU_KLEAE	P12381 klebsiella
28	9	22.0	9	UN19_CLOPA	P81355 clostridium
29	8	19.5	4	FYRI_ATEL	P58706 anthopleura
30	8	19.5	5	SUGA_ACHDO	P19991 acheta dome
31	8	19.5	7	MNP1_LEPDE	P42984 leptinotars
32	8	19.5	7	WNA1_ACHFU	P35919 achatina fu
33	8	19.5	7	WNA3_ACHFU	P35921 achatina fu

#### ALIGNMENTS

##### RESULT 1

UH09\_RAT ID UH09\_RAT STANDARD; PRT; 8 AA.  
AC P56575;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE Unknown protein from 2D-page of heart tissue (Spot P9) (Fragment).  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=Wistar; TISSUE=Heart;  
RA Li X.-P., Pleissner K.-P., Scheler C., Regitz-Zagrosek V., Salikov J.,  
RA Jungblut P.R.;  
RL Submitted (SEP-1998) to the SWISS-PROT data bank.  
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
CC PROTEIN IS: 8.9, ITS MW IS: 42 kDa.  
FT NON\_TER 8  
SQ SEQUENCE 8 AA; 1029 MW; 9E0775A6C4140B06 CRC64;

Query Match 48.8%; Score 20; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 QERR 8  
DB 1 QERR 4

##### RESULT 2

ULAK\_MOUSE ID ULAK\_MOUSE STANDARD; PRT; 9 AA.  
AC P99031;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE Unknown protein from 2D-page of liver tissue (Spot 2D-0014LD) (Fragment).  
DE Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Liver;  
RA Sanchez J.-C., Rouge V., Frutiger S., Hughes G.J., Yan J.X.,  
RA Hoogland C., Appel R.D., Binz P.-A., Hochstrasser D.F.,  
RA Cowthorne M.;  
RL Submitted (AUG-1998) to the SWISS-PROT data bank.  
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
CC PROTEIN IS: 6.0, ITS MW IS: 12.5 kDa.  
DR SWISS-2DPAGE; P99031; MOUSE.  
FT NON\_TER 9

34 8 19.5 8 1 ALI2\_CARMA P81815 carcinus ma  
35 8 19.5 9 1 ALI10\_CARMA P81813 carcinus ma  
36 8 19.5 9 1 DSIP\_RABIT P01158 oryctolagus  
37 8 19.5 9 1 FAR2\_CALVO P41857 calliphora  
38 8 19.5 9 1 FAR3\_CALVO P41858 calliphora  
39 8 19.5 9 1 FAR6\_CALVO P41861 calliphora  
40 8 19.5 9 1 FPF1\_SARBU P83350 sarcophaga  
41 8 19.5 9 1 IPYR\_RHOVI P82992 rhodopseuso  
42 8 19.5 9 1 LMIP\_LOCOMI P31799 locusta mig  
43 8 19.5 9 1 PTSP\_BOMMO P82003 bombyx mori  
44 8 19.5 9 1 RT33\_BOVIN P82926 bos taurus  
45 7 17.1 4 1 DCWS\_PSECH P19918 pseudomonas

SQ SEQUENCE 9 AA; 1106 MW; E1E842C3240B145A CRC64;

Query Match 39.0%; Score 16; DB 1; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 1.3e+05;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 ERRV 9  
 ||:|  
 Db 4 ERKV 7

RESULT 3  
 ID ACPH\_RABIT STANDARD; PRT; 6 AA.  
 AC P25154;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Acylamino-acid-releasing enzyme (EC 3.4.19.1) (Acyl-peptide hydrolase)  
 DE (APH) (Acylaminoacyl-peptidase) (Fragment).  
 GN APEH.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Muscle;  
 RX MEDLINE=92222120; PubMed=1807161;  
 RA Krishna R.G., Chin C.C.O., Wolf F.;  
 RT "N-terminal sequence analysis of N alpha-acetylated proteins after  
 unblocking with N-acylaminoacyl-peptide hydrolase.";  
 RL Anal. Biochem. 199;45:50(1991).  
 CC -1- FUNCTION: THIS ENZYME CATALYZES THE HYDROLYSIS OF THE AMINO-  
 TERMINAL PEPTIDE BOND OF AN N-ACETYLATED PEPTIDE TO GENERATE  
 AN N-ACETYLATED AA AND A PEPTIDE WITH A FREE AMINO-TERMINUS.  
 CC IT PREFERENTIALLY CLEAVES OFF AC-ALA, AC-MET AND AC-SER.  
 CC -1- CATALYTIC ACTIVITY: Acylaminoacyl-peptide + H(2)O -> acylamino acid  
 + peptide.  
 CC -1- SUBUNIT: Homotetramer.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S9C.  
 CC PIR: A49792; A49792.  
 DR MEROPS: S09.004.-;  
 DR InterPro: IPR002471; Prol\_endopep\_ser.  
 DR PROSITE: PS00708; PRO-ENDOPEP\_SER; PARTIAL.  
 KW Hydrolase; Acetylation.  
 FT MOD\_RES 1 1 ACETYLATION.  
 FT NON\_TER 6 6  
 SQ SEQUENCE 6 AA; 775 MW; 6732D6C40B16F000 CRC64;

Query Match 36.6%; Score 15; DB 1; Length 6;  
 Best Local Similarity 75.0%; Pred. No. 1.3e+05;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 ERRV 9  
 ||:|  
 Db 2 ERQV 5

RESULT 4  
 ID B44K\_PORGI STANDARD; PRT; 8 AA.  
 AC P8186;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE 44 kDa immunogenic protein (Fragment).  
 OS Porphyromonas gingivalis (Bacteroides gingivalis).  
 OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;  
 OC Porphyromonadaceae; Porphyromonas.  
 OX NCBI\_TaxID=837;  
 RN [1]  
 RP SEQUENCE.

RC STRAIN=VPB 3492;  
 RX MEDLINE=20198497; PubMed=10731616;  
 RA Norris J.M., Love D.N.;  
 RT "Serum antibody responses of cats to soluble whole cell antigens of  
 feline Porphyromonas gingivalis.";  
 RL Vet. Microbiol. 73:37-49(2000).  
 CC -1- SIMILARITY: TO P.GINGIVALIS HEMAGGLUTININ A.  
 KW Antigen.  
 FT NON\_TER 8 8  
 SQ SEQUENCE 8 AA; 989 MW; 9554540326CB476D CRC64;

Query Match 34.1%; Score 14; DB 1; Length 8;  
 Best Local Similarity 40.0%; Pred. No. 1.3e+05;  
 Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 QERV 9  
 ||:|  
 Db 4 QKRNI 8

RESULT 5  
 ID CLP\_THICU STANDARD; PRT; 8 AA.  
 AC P80488;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Chemolithotroph-specific protein (Fragment).  
 OS Thiobacillus cuprinus.  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Comamonadaceae; Thiomonas.  
 OX NCBI\_TaxID=36860;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=DSM 5494;  
 RA Marin I., Anaro A.M., Jerez C.A., Amils R., Abad J.P.;  
 RL Submitted (SEP-1995) to the SWISS-PROT data bank.  
 CC -1- MISCELLANEOUS: FOUND SPECIFICALLY IN CELLS CULTURED  
 CHEMOLITHOTROPHICALLY.  
 FT NON\_TER 8 8  
 SQ SEQUENCE 8 AA; 785 MW; 91487B06DDC2D76D CRC64;

Query Match 34.1%; Score 14; DB 1; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AQE 6  
 |||  
 Db 4 AQE 6

RESULT 6  
 ID RS10\_SERMA STANDARD; PRT; 9 AA.  
 AC O68936;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE 30S ribosomal protein S10 (Fragment).  
 GN RPSJ.  
 OS Serratia marcescens.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Serratia.  
 OX NCBI\_TaxID=615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Noorani S.M., Lindahl L., Zengel J.M.;  
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Involved in the binding of tRNA to the ribosomes (By  
 similarity).  
 CC -1- SIMILARITY: BELONGS TO THE S10P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL: AF058451; AAC14294.1; -;  
 CC HAMAP; MF\_00508; -; 1.  
 CC InterPro; IPR001848; Ribosomal\_S10.  
 CC PROSITE; PS00361; RIBOSOMAL\_S10; PARTIAL.  
 CC Ribosomal protein.  
 CC NON\_TER 9  
 CC SEQUENCE 9 AA; 1214 MW; DE3944004416D456 CRC64;

Query Match 34.1%; Score 14; DB 1; Length 9;  
 Best Local Similarity 40.0%; Pred. No. 1.3e+05;  
 Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 5 QRRV 9  
 DB 2 QNQRI 6

## RESULT 7

UPAL\_HUMAN STANDARD; PRT; 8 AA.  
 AC P30087;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Unknown protein from 20-page of plasma (Spot 2) (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.

RC TISSUE=Plasma;  
 RX MEDLINE=93092937; PubMed=1459097;  
 RA Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,  
 RA Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,  
 RA Hochstrasser D.F.;  
 RT "Plasma protein map: an update by microsequencing."  
 RL Electrophoresis 13:707-714(1992).  
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 CC PROTEIN IS: 4.9, ITS MW IS: 65 kDa.  
 DR SWISS-2DPAGE; P30087; HUMAN.  
 FT NON\_TER 1  
 FT UNSURE 8  
 FT NON\_TER 8  
 SQ SEQUENCE 8 AA; 944 MW; C01772C455BB06DA CRC64;

Query Match 31.7%; Score 13; DB 1; Length 8;  
 Best Local Similarity 60.0%; Pred. No. 1.3e+05;  
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 5 QRRV 9  
 DB 2 QESNV 6

## RESULT 8

UH11\_RAT STANDARD; PRT; 7 AA.  
 AC P56576;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Unknown protein from 2D-page of heart tissue (Spot P11) (Fragment).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE.

RC STRAIN=Wistar; TISSUE=Heart;  
 RA Li X.-P., Pleissner K.-P., Scheler C., Regitz-Zagrosek V., Salikov J.,  
 RA Jungblut P.R.;  
 RL Submitted (SEP-1998) to the SWISS-PROT data bank.  
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 CC PROTEIN IS: 8.5, ITS MW IS: 42 kDa.  
 CC UNSURE 2  
 CC NON\_TER 7  
 CC SEQUENCE 7 AA; 775 MW; 6866DB040DC5A6B0 CRC64;

Query Match 29.3%; Score 12; DB 1; Length 7;  
 Best Local Similarity 40.0%; Pred. No. 1.3e+05;  
 Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 3 AAOER 7  
 DB 2 SAREQ 6

## RESULT 9

NS3\_MYCTU STANDARD; PRT; 8 AA.  
 AC P81152;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE 30 kDa non-secretory protein 3 (Fragment).  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1773;  
 RN [1]  
 RP SEQUENCE.

RC STRAIN=H37RV;  
 RA Prasad H.K., Annappurna P.S.;  
 RL Submitted (DEC-1997) to the SWISS-PROT data bank.  
 CC -!- CAUTION: We are unable to find this protein in the translation of  
 CC the genome of strain H37RV.  
 FT NON\_TER 1  
 FT NON\_TER 8  
 SQ SEQUENCE 8 AA; 919 MW; 8D3DC40B19CDC2D2 CRC64;

Query Match 29.3%; Score 12; DB 1; Length 8;  
 Best Local Similarity 50.0%; Pred. No. 1.3e+05;  
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 LAAQER 7  
 DB 1 VVAER 6

## RESULT 10

FARD\_CALVO STANDARD; PRT; 9 AA.  
 AC P41868;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE Calliphras famide 13.  
 OS Calliphras vomitoria (Blue blowfly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;  
 OC Calliphoridae; Calliphora.  
 OX NCBI\_TaxID=27454;  
 RN [1]  
 RP SEQUENCE.

RC TISSUE=thoracic ganglion;  
 RX MEDLINE=92196111; PubMed=1549595;  
 RA Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,  
 RA Rehfeld J.F., Thorpe A.;

```

RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2
RT neuropeptides (designated calliPMPFamides) from the blowfly
RT Calliphora vomitoria.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).
CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
DR PIR: D44787; D44787.
KW Neuropeptide; Amidation.
FT MOD_RES 9 9
SQ SEQUENCE 9 AA; 1028 MW; 22D10699C87AB6D8 CRC64;

Query Match 26.8%; Score 11; DB 1; Length 9;
Best Local Similarity 50.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 AAQE 6
DB 1 AQD 4

RESULT 11
ULAH_HUMAN STANDARD; PRT; 9 AA.
ID P31934;
AC P31934;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Unknown protein from 2D-page of liver tissue (Spot 94) (Fragment).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=94147969; PubMed=8313870;
RA Hughes G.J., Frutiger S., Paquet N., Pasquali C., Sanchez J.-C.,
RA Tissot J.-D., Balroch A., Appel R.D., Hochstrasser D.F.;
RT "Human liver protein map: update 1993.";
RL Electrophoresis 14:1216-1222(1993).
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 5.5, ITS MW IS: 12 kDa.
DR SWISS-2DPAGE; P31934; HUMAN.
FT NON_TER 1 1
FT VARIANT 3 3
FT FT A -> L.
FT FT /FTid=VAR_000001.
SQ SEQUENCE 9 AA; 990 MW; FICC7B172AB6DDCA CRC64;

Query Match 26.8%; Score 11; DB 1; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.3e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 AQE 6
DB 3 AQD 5

RESULT 12
CIA_ENTFA STANDARD; PRT; 7 AA.
ID P11932;
AC P11932;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Sex pheromone cAM373 (Clumping-inducing agent) (CIA).
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Lactobacillales; Enterococaceae; Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RP SEQUENCE.
RX MEDLINE=87005252; PubMed=3093276;
RA Mori M., Tanaka H., Sakagami Y., Isogai A., Fujino M., Kitada C.,

```

```

RA White B.A., An F.Y., Clewell D.B., Suzuki A.;
RT "Isolation and structure of the Streptococcus faecalis sex pheromone,
RT cAM373.";
RL FEBS Lett. 206:69-72(1986).
CC -1- FUNCTION: CAM373 INDUCES MATING RESPONSE OF DONOR CELLS
CC HARBORING PAM373.
CC -1- MISCELLANEOUS: THE N-TERMINUS IS POSSIBLY RESPONSIBLE FOR
CC SPECIFICITY OF PHEROMONES TO PLASMIDS.
CC -1- SIMILARITY: C-TERMINAL TO THE PHEROMONES CPD1 AND CAD1.
DR PIR: A25269; A25269.
KW Pheromone.
SQ SEQUENCE 7 AA; 734 MW; 75BDD72059C05DB0 CRC64;

Query Match 24.4%; Score 10; DB 1; Length 7;
Best Local Similarity 50.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAA 4
DB 4 ILAS 7

RESULT 13
NPB_BOVIN STANDARD; PRT; 8 AA.
ID P15507;
AC P15507;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Neuropeptide B.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=86067985; PubMed=3865193;
RA Yang H.-Y.T., Fratta W., Majane E.A., Costa E.;
RT "Isolation, sequencing, synthesis, and pharmacological
RT characterization of two brain neuropeptides that modulate the action
RT of morphine.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:7757-7761(1985).
CC -1- FUNCTION: MODULATES THE ACTION OF MORPHINE.
DR PIR: B24749; B24749.
KW Neuropeptide; Amidation.
FT MOD_RES 8 8
SQ SEQUENCE 8 AA; 1082 MW; 87D416C776D9C729 CRC64;

Query Match 24.4%; Score 10; DB 1; Length 8;
Best Local Similarity 50.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 QERR 8
DB 4 QPOR 7

RESULT 14
UC26_MAIZE STANDARD; PRT; 8 AA.
ID UC26_MAIZE
AC P80632;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Unknown protein from 2D-page of etiolated coleoptile (Spot 907)
DE (fragment).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;

```



Job time : 14 secs

```
RN [1]
RP SEQUENCE.
RC TISSUE=Coleoptile;
RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
RA Pernollet J.-C., Zivy M., de Vienne D.;
RT "The maize two dimensional gel protein database: towards an integrated
RT genome analysis program.";
RL Theor. Appl. Genet. 93:997-1005(1996).
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 7.0, ITS MW IS: 57.2 kDa.
DR Maize-2DPAGE; P80632; COLEOPTILE.
FT NON_TER 1 1
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 990 MW; 9639D6DAB4176B1D CRC64;

Query Match 24.4%; Score 10; DB 1; Length 8;
Best Local Similarity 50.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 AQER 7
DB 1 AEPR 4

RESULT 15
FIBB_PAPAN
ID FIBB_PAPAN STANDARD; PRT; 9 AA.
AC P19344;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).
GN FGB.
OS Papio anubis (Olive baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Papio.
OX NCBI_TaxID=9555;
RN [1]
RP SEQUENCE.
RX MEDLINE=84161822; PubMed=6423621;
RA Nakamura S., Takenaka O., Takahashi K.;
RT "Fibrinopeptides A and B of baboons (Papio anubis, Papio hamadryas,
RT and Theropithecus gelada): their amino acid sequences and
RT evolutionary rates and a molecular phylogeny for the baboons.";
RL J. Biochem. 94:1973-1978(1983).
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC AGGREGATION.
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
DR PIR; D28854; D28854.
DR InterPro; IPR002181; Fibrinogen_C.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
KW Blood coagulation; Plasma.
FT PEPTIDE 1 9 FIBRINOPEPTIDE B.
FT NON_TER 9 9
SQ SEQUENCE 9 AA; 1076 MW; DDF6409C7287B06 CRC64;

Query Match 24.4%; Score 10; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 QE 6
DB 2 QE 3
```

Search completed: October 7, 2003, 14:46:06

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:41:36 ; Search time 58.5 Seconds  
(without alignments)  
39.700 Million cell updates/sec

Title: US-09-807-512-26

Perfect score: 41

Sequence: 1 MLAGERRV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 775

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_23.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rhodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	17	41.5	9	15 Q85710	Q85710 rous sarcom
2	16	39.0	7	2 Q47477	Q47477 escherichia
3	15	36.6	9	4 Q15892	Q15892 homo sapien
4	15	36.6	9	10 Q940K4	Q940K4 arabidopsis
5	15	36.6	9	11 Q88889	Q88889 mus musculu
6	14	34.1	8	5 Q15896	Q15896 babesia bov
7	14	34.1	9	2 Q9K4M6	Q9K4M6 staphylococ
8	14	34.1	9	6 Q9TT77	Q9TT77 bos taurus
9	14	34.1	9	6 Q9TRW2	Q9TRW2 oryctolagus
10	13	31.7	7	8 P92214	P92214 amblyopyrum
11	13	31.7	7	8 P92393	P92393 hordeum vul
12	13	31.7	7	8 P92403	P92403 lophopyrum
13	13	31.7	7	8 P92427	P92427 peridictyon
14	13	31.7	7	8 P92430	P92430 aegilops ta
15	13	31.7	7	8 P92221	P92221 bromus iner
16	13	31.7	7	8 P92425	P92425 pseudoroegn

17	13	31.7	7	8 P92381	P92381 hordeum bra
18	13	31.7	7	8 P92387	P92387 henrardia p
19	13	31.7	7	8 P92210	P92210 agropyron c
20	13	31.7	7	8 P92440	P92440 thiropyrum
21	13	31.7	7	8 P92218	P92218 australopyr
22	13	31.7	7	8 P92390	P92390 heteranthe
23	13	31.7	7	8 P92372	P92372 haynaldia v
24	13	31.7	7	8 P92442	P92442 taeniatheru
25	13	31.7	7	8 P92226	P92226 crithopsis
26	13	31.7	7	8 P92385	P92385 hordeum mar
27	13	31.7	7	8 P92421	P92421 psathyrosta
28	13	31.7	8	2 Q32560	Q32560 escherichia
29	13	31.7	8	2 Q8RSR3	Q8RSR3 lactobacill
30	13	31.7	8	11 Q62528	Q62528 mus spretus
31	13	31.7	9	2 Q53914	Q53914 streptomyce
32	13	31.7	9	12 Q66545	Q66545 human herpe
33	12	29.3	8	4 Q9Y4J4	Q9Y4J4 homo sapien
34	12	29.3	8	11 Q99MN0	Q99MN0 mus musculu
35	12	29.3	8	12 Q83977	Q83977 influenzavi
36	12	29.3	9	2 Q45852	Q45852 clostridium
37	12	29.3	9	2 Q46179	Q46179 clostridium
38	12	29.3	9	2 Q93LE4	Q93LE4 heliobacill
39	12	29.3	9	2 Q9R5M1	Q9R5M1 staphylococ
40	12	29.3	9	12 Q83622	Q83622 murray vall
41	12	29.3	9	15 Q8UTD7	Q8UTD7 human immun
42	11	26.8	7	12 Q9YVE3	Q9YVE3 human adeno
43	11	26.8	7	12 Q9YIO9	Q9YIO9 human adeno
44	11	26.8	7	12 Q9YIRO	Q9YIRO human adeno
45	11	26.8	8	2 Q9R7T2	Q9R7T2 escherichia

ALIGNMENTS

RESULT 1

Q85710 ID Q85710 PRELIMINARY; PRT; 9 AA.  
AC Q85710;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
DE Pol protein (Fragment).  
OS Rous sarcoma virus.  
OC Viruses; Retroid viruses; Retroviridae; Alpharetrovirus.  
OX NCBI\_TaxID=11886;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=84115080; PubMed=6319754;  
RA Lerner T.L., Hanafusa H.;  
RT "DNA sequence of the Bryan high-titer strain of Rous sarcoma virus:  
RT Extent of env deletion and possible genealogical relationship with  
RT other viral strains".  
RL J. Virol. 49:549-556(1984).  
DR EMBL; K03365; AAA42557.1; -.  
FT NON\_TER 1  
SQ SEQUENCE 9 AA; 949 MW; 94AA144DDDD731AA CRC64;

Query Match 41.5%; Score 17; DB 15; Length 9;  
Best Local Similarity 80.0%; Pred. No. 8.3e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAAQ 6  
|||  
Db 4 LAANE 8

RESULT 2

Q47477 ID Q47477 PRELIMINARY; PRT; 7 AA.  
AC Q47477;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)

```

DE Tpi protein (Fragment).
GN TPI.
OS Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE OF 7-7 FROM N.A.
RX MEDLINE=85203917; PubMed=3158524;
RA Hellinga H.W., Evans P.R.;
RT "Nucleotide sequence and high-level expression of the major
RT Escherichia coli phosphofructokinase.";
RL Eur. J. Biochem. 149:363-373(1985).
RN [2]
RP SEQUENCE FROM N.A.
RA Evans P.;
RL Submitted (OCT-1986) to the EMBL/GenBank/DBJ databases.
DR EMBL; X02519; CAA26359.1; -.
FT NON_TER 1
SQ SEQUENCE 7 AA; 773 MW; 7416D33DDDB1DB0 CRC64;

Query Match 39.0%; Score 16; DB 2; Length 7;
Best Local Similarity 60.0%; Pred. No. 8.3e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAQQR 7
Db 3 AAQQR 7

RESULT 3
Q15892
ID Q15892 PRELIMINARY; PRT; 9 AA.
AC Q15892;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE (Clone XP3B4A) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Lee C.-C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,
RA Coolbaugh M.I., Chinault C.A., Baldini A., Lindsay E.A., Zhao Z.-Y.,
RA Caskey C.T.H.;
RT "Isolation of chromosome-specific genes by reciprocal probing of
RT arrayed cDNAs and cosmid libraries.";
RL Hum. Mol. Genet. 0:0-0(1995).
DR EMBL; L32071; AAA73882.1; -.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 9 AA; 971 MW; 49B22732CDC40B17 CRC64;

Query Match 36.6%; Score 15; DB 4; Length 9;
Best Local Similarity 66.7%; Pred. No. 8.3e+05;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4'AQERRV 9
Db 1 ALERAV 6

RESULT 4
Q940K4
ID Q940K4 PRELIMINARY; PRT; 9 AA.
AC Q940K4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 1.4 kDa protein.

```

```

OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsals.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY054502; AAK96693.1; -.
DR EMBL; AY081672; AAM10234.1; -.
KW Hypothetical protein.
SQ SEQUENCE 9 AA; 1362 MW; 712D1416D9D41414 CRC64;

Query Match 36.6%; Score 15; DB 10; Length 9;
Best Local Similarity 75.0%; Pred. No. 8.3e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 QERR 8
Db 6 QRRR 9

RESULT 5
O88889
ID O88889 PRELIMINARY; PRT; 9 AA.
AC O88889;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Ubiquitin-conjugating enzyme UBCM4 (Fragment).
GN UBCM4
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RX MEDLINE=97057256; PubMed=8901595;
RA Harbers K., Muller U., Grams A., Li E., Jaenisch R., Franz T.;
RT "Provirus integration into a gene encoding a ubiquitin-conjugating
RT enzyme results in a placental defect and embryonic lethality.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:12412-12417(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RX MEDLINE=99132641;
RA Miller U., Grams A., Martinez-Noel G., Copeland N.G., Gilbert D.J.,
RA Jenkins N.A., Harbers K.;
RT "Structure of the gene encoding the ubiquitin-conjugating enzyme
RT UbcM4, characterization of its promoter, and chromosomal location.";
RL Gene 224:109-116(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RA Mueller U., Grams A., Martinez-Noel G., Harbers K.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF071557; AAD10128.1; -.

```

```
FT NON_TER 9 9
SQ SEQUENCE 9 AA; 1063 MW; C90F97341415BDD CRC64;

Query Match 36.6%; Score 15; DB 11; Length 9;
Best Local Similarity 50.0%; Pred. No. 8.3e+05;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LAAQER 7
   :|||
Db 1 MAASRR 6

RESULT 6
O15896 ID O15896 PRELIMINARY; PRT; 8 AA.
AC O15896;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE 12D3 antigen (Fragment).
GN 12D3.
OS Babesia bovis.
OC Eukaryota; Alveolata; Apicomplexa; Piroplasmida; Babesiidae; Babesia.
OX NCBI_TaxID=5855;
RN [1]
RC STRAIN=Samford attenuated;
RA Silins G.U., Blakeley R.L., Riddles P.W.;
RT "Characterization of a gene encoding an ATP-binding protein from Babesia bovis.";
RL Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
RN [2]
RC SEQUENCE FROM N.A.
RP STRAIN=Samford attenuated;
RA Silins G.U., Blakeley R.L., Riddles P.W.;
RT "Characterization of the transcriptional control region of the 12D3 antigen gene from the sporozoan Babesia bovis.";
RL Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
RN [3]
RC SEQUENCE FROM N.A.
RP Silins G., Blakeley R., Riddles P.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; U4917; AAB66362.1; -.
DR EMBL; AY10919; AAN64587.1; -.
FT NON_TER 8
SQ SEQUENCE 8 AA; 984 MW; F0D2C9D411ADD726 CRC64;

Query Match 34.1%; Score 14; DB 5; Length 8;
Best Local Similarity 60.0%; Pred. No. 8.3e+05;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAQ 5
   :|||
Db 1 MLATR 5

RESULT 7
O9K4M6 ID O9K4M6 PRELIMINARY; PRT; 9 AA.
AC O9K4M6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Transposase IS257 (Fragment).
GN TNPA.
OS Staphylococcus haemolyticus.
OG Plasmid pNVH97A.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1283;
RN [1]
RC SEQUENCE FROM N.A.
RA Anthonisen I., Steinum T.M., Sidhu M.S., Sorum H., Sunde M.;
```

```
RT "Organization of the beta-lactamase transposon Tn552, the antiseptic resistant gene qacA and a Staphylococcal insertion Sequence IS257 on large plasmids in multiresistant Staphylococcus haemolyticus.";
RL Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AJ400722; CAB94812.1; -.
KW Plasmid.
FT NON_TER 1
SQ SEQUENCE 9 AA; 1000 MW; 6FD5D7269045B05B CRC64;

Query Match 34.1%; Score 14; DB 2; Length 9;
Best Local Similarity 75.0%; Pred. No. 8.3e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAA 4
   :|||
Db 6 MLAS 9

RESULT 8
Q9TT77 ID Q9TT77 PRELIMINARY; PRT; 9 AA.
AC Q9TT77;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Wilm's tumor protein 1 (Fragment).
GN WT1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RC SEQUENCE FROM N.A.
RP MEDLINE=21015404; PubMed=11130975;
RA Brouillette J.A., Andrew J.R., Venta P.J.;
RT "Estimate of nucleotide diversity in dogs with a pool-and-sequence method.";
RL Mamm. Genome 11:1079-1086(2000).
DR EMBL; AF202074; AAF20919.1; -.
FT NON_TER 9
SQ SEQUENCE 9 AA; 1231 MW; 58DDF41416D1F403 CRC64;

Query Match 34.1%; Score 14; DB 6; Length 9;
Best Local Similarity 42.9%; Pred. No. 8.3e+05;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 LAAQERR 8
   :|||
Db 1 LKRHQR 7

RESULT 9
Q9TRW2 ID Q9TRW2 PRELIMINARY; PRT; 9 AA.
AC Q9TRW2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CALDESMON=PHOSPHORYLATION site (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RC SEQUENCE.
RP MEDLINE=91378498; PubMed=1898046;
RA Ikebe M., Hornick T.;
RT "Determination of the phosphorylation sites of smooth muscle caldesmon by protein kinase C.";
RL Arch. Biochem. Biophys. 288:538-542(1991).
FT NON_TER 1
```

FT NON\_TER 9 9 1018 MW; 8C901B10533735A5 CRC64;  
SQ SEQUENCE 9 AA; 1018 MW; 8C901B10533735A5 CRC64;

Query Match 34.1%; Score 14; DB 6; Length 9;  
Best Local Similarity 50.0%; Pred. No. 8.3e+05;  
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 LAAQER 7  
Db 4 LKIEER 9

## RESULT 10

P92214  
ID P92214 PRELIMINARY; PRT; 7 AA.  
AC P92214; 1997 (TrEMBLrel. 03, Created)  
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Ribosomal protein 11 (Fragment).  
GN RPS11.  
OS Amblyopyrum muticum.  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
OC Triticeae; Amblyopyrum.  
OX NCBI\_TaxID=4595;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=H5572; TISSUE=Leaf;  
RX MEDLINE=97271648; PubMed=9126564;  
RA Petersen G., Seberg O.;  
RT "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA  
sequence data."; Evol. 7:217-230(1997).  
RL Mol. Phylogenet. Evol. 7:217-230(1997).  
DR EMBL; Z77756; CAB01346.1; -.  
KW Chloroplast.  
FT NON\_TER 1  
SQ SEQUENCE 7 AA; 894 MW; 6734140333277700 CRC64;

Query Match 31.7%; Score 13; DB 8; Length 7;  
Best Local Similarity 40.0%; Pred. No. 8.3e+05;  
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 5 QERRV 9  
Db 3 KKRRRL 7

## RESULT 11

P92393  
ID P92393 PRELIMINARY; PRT; 7 AA.  
AC P92393;  
DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Ribosomal protein 11 (Fragment).  
GN RPS11.  
OS Hordeum vulgare (Barley).  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
OC Triticeae; Hordeum.  
OX NCBI\_TaxID=4513;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=H3139; TISSUE=Leaf;  
RX MEDLINE=97271648; PubMed=9126564;  
RA Petersen G., Seberg O.;  
RT "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA  
sequence data."; Evol. 7:217-230(1997).  
RL Mol. Phylogenet. Evol. 7:217-230(1997).  
DR EMBL; Z77764; CAB01370.1; -.

KW Chloroplast.  
FT NON\_TER 1  
SQ SEQUENCE 7 AA; 894 MW; 6734140333277700 CRC64;

Query Match 31.7%; Score 13; DB 8; Length 7;  
Best Local Similarity 40.0%; Pred. No. 8.3e+05;  
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 5 QERRV 9  
Db 3 KKRRRL 7

## RESULT 12

P92403  
ID P92403 PRELIMINARY; PRT; 7 AA.  
AC P92403;  
DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Ribosomal protein 11 (Fragment).  
GN RPS11.  
OS Lophopyrum elongatum (Tall wheatgrass) (Argopyrum elongatum).  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
OC Triticeae; Lophopyrum.  
OX NCBI\_TaxID=4588;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=H6692; TISSUE=Leaf;  
RX MEDLINE=97271648; PubMed=9126564;  
RA Petersen G., Seberg O.;  
RT "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA  
sequence data."; Evol. 7:217-230(1997).  
RL Mol. Phylogenet. Evol. 7:217-230(1997).  
DR EMBL; Z77743; CAB01307.1; -.  
KW Chloroplast.  
FT NON\_TER 1  
SQ SEQUENCE 7 AA; 894 MW; 6734140333277700 CRC64;

Query Match 31.7%; Score 13; DB 8; Length 7;  
Best Local Similarity 40.0%; Pred. No. 8.3e+05;  
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 5 QERRV 9  
Db 3 KKRRRL 7

## RESULT 13

P92427  
ID P92427 PRELIMINARY; PRT; 7 AA.  
AC P92427;  
DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Ribosomal protein 11 (Fragment).  
GN RPS11.  
OS Peridictyon sanctum.  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
OC Triticeae; Peridictyon.  
OX NCBI\_TaxID=37683;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=H5575; TISSUE=Leaf;  
RX MEDLINE=97271648; PubMed=9126564;  
RA Petersen G., Seberg O.;  
RT "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA  
sequence data."; Evol. 7:217-230(1997).  
RL Mol. Phylogenet. Evol. 7:217-230(1997).

DR EMBL; 277749; CAB01325.1; -  
KW Chloroplast.  
FT NON\_TER 1  
SQ SEQUENCE 7 AA; 894 MW; 6734140333277700 CRC64;

Query Match 31.7%; Score 13; DB 8; Length 7;  
Best Local Similarity 40.0%; Pred. No. 8.3e+05;  
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 5 QERRV 9  
::||:  
Db 3 KKRL 7

## RESULT 14

P92430  
ID P92430 PRELIMINARY; PRT; 7 AA.  
AC P92430;  
DT 01-MAY-1997 (TREMBLrel. 03, Created)  
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Ribosomal protein 11 (Fragment).  
GN rps11.  
OS Aegilops tauschii (Aegilops squarrosa).  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;  
OC Triticeae; Aegilops.  
OX NCBI\_TaxID=37682;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=H6668; TISSUE=Leaf;  
RX MEDLINE=97271648; PubMed=9126564;  
RA Petersen G.; Seberg O.;  
RT "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA  
sequence data."  
RL Mol. Phylogenet. Evol. 7:217-230(1997).  
DR EMBL; 277758; CAB01352.1; -  
KW Chloroplast.  
FT NON\_TER 1  
SQ SEQUENCE 7 AA; 894 MW; 6734140333277700 CRC64;

Query Match 31.7%; Score 13; DB 8; Length 7;  
Best Local Similarity 40.0%; Pred. No. 8.3e+05;  
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 5 QERRV 9  
::||:  
Db 3 KKRL 7

## RESULT 15

P92221  
ID P92221 PRELIMINARY; PRT; 7 AA.  
AC P92221;  
DT 01-MAY-1997 (TREMBLrel. 03, Created)  
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Ribosomal protein 11 (Fragment).  
GN rps11.  
OS Bromus inermis (Smooth brome grass).  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;  
OC Bromaeae; Bromus.  
OX NCBI\_TaxID=15371;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=OSA414; TISSUE=Leaf;  
RX MEDLINE=97271648; PubMed=9126564;  
RA Petersen G.; Seberg O.;  
RT "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA  
sequence data."

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:43:36 ; Search time 20 Seconds  
(without alignments)  
19.040 Million cell updates/sec

Title: US-09-807-512-26

Perfect score: 41

Sequence: 1 MLAGERRV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 77717

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	61.0	9	3 US-08-144-779C-27	Sequence 27, Appl
2	25	61.0	9	4 US-09-368-449B-27	Sequence 27, Appl
3	22	53.7	7	3 US-08-144-779C-24	Sequence 24, Appl
4	22	53.7	7	4 US-09-368-449B-24	Sequence 24, Appl
5	22	53.7	8	3 US-08-144-779C-25	Sequence 25, Appl
6	22	53.7	8	4 US-09-368-449B-25	Sequence 25, Appl
7	22	53.7	9	3 US-08-144-779C-26	Sequence 26, Appl
8	22	53.7	9	4 US-09-368-449B-26	Sequence 26, Appl
9	22	53.7	9	4 US-09-447-399-7	Sequence 7, Appl
10	21	51.2	6	4 US-08-055-006-53	Sequence 53, Appl
11	21	51.2	7	4 US-09-627-851B-39	Sequence 39, Appl
12	21	51.2	9	3 US-09-518-046-78	Sequence 78, Appl
13	21	51.2	9	3 US-09-518-046-95	Sequence 95, Appl
14	21	51.2	9	3 US-09-518-046-146	Sequence 146, Appl
15	21	51.2	9	4 US-09-644-600-59	Sequence 59, Appl
16	20	48.8	6	3 US-08-463-160B-65	Sequence 65, Appl
17	20	48.8	7	1 US-08-552-142A-5	Sequence 5, Appl
18	20	48.8	7	1 US-08-910-973-5	Sequence 5, Appl
19	20	48.8	7	4 US-09-499-227-5	Sequence 5, Appl
20	20	48.8	7	5 PCT-US95-05741-5	Sequence 31, Appl
21	20	48.8	8	3 US-08-144-779C-31	Sequence 31, Appl
22	20	48.8	8	4 US-09-368-449B-31	Sequence 28, Appl
23	20	48.8	9	4 US-09-434-476A-28	Sequence 28, Appl
24	20	48.8	9	4 US-09-434-476A-29	Sequence 29, Appl
25	19	46.3	5	4 US-09-638-202A-63	Sequence 63, Appl
26	19	46.3	6	2 US-08-318-837-22	Sequence 22, Appl
27	18	43.9	6	2 US-08-446-345-17	Sequence 17, Appl

28	18	43.9	7	3 US-09-413-452-48	Sequence 48, Appl
29	18	43.9	7	3 US-09-413-068-48	Sequence 48, Appl
30	18	43.9	9	1 US-08-189-772-4	Sequence 4, Appl
31	18	43.9	9	1 US-08-188-277B-13	Sequence 13, Appl
32	18	43.9	9	3 US-09-510-738A-32	Sequence 32, Appl
33	18	43.9	9	3 US-09-510-738A-61	Sequence 61, Appl
34	18	43.9	9	3 US-09-510-738A-171	Sequence 171, Appl
35	18	43.9	9	4 US-09-461-325-260	Sequence 260, Appl
36	18	43.9	9	4 US-09-434-476A-30	Sequence 30, Appl
37	18	43.9	9	4 US-09-861-966-32	Sequence 32, Appl
38	18	43.9	9	4 US-09-861-966-61	Sequence 61, Appl
39	18	43.9	9	4 US-09-861-966-171	Sequence 171, Appl
40	18	43.9	9	4 US-08-983-157B-16	Sequence 16, Appl
41	17	41.5	5	5 PCT-US93-12679-39	Sequence 39, Appl
42	17	41.5	6	1 US-08-121-202-8	Sequence 8, Appl
43	17	41.5	6	3 US-08-144-779C-28	Sequence 28, Appl
44	17	41.5	6	3 US-08-888-381-11	Sequence 11, Appl
45	17	41.5	6	4 US-09-368-449B-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1  
US-08-144-779C-27  
; Sequence 27, Application US/08144779C  
; Patent No. 6066622  
; GENERAL INFORMATION:  
; APPLICANT: Green, Lawrence R.  
; APPLICANT: Snackevich, Nicolay V.  
; APPLICANT: Ivanov, Vadim T.  
; APPLICANT: Mikhailoyova, Inessa I.  
; APPLICANT: Vaskovsky, Boris V.  
; APPLICANT: Mikhailov, Alexander N.  
; APPLICANT: Khavinson, Vladimir K.  
; APPLICANT: Morozov, Vyacheslav G.  
; TITLE OF INVENTION: Immunomodulating Peptides and Methods  
; TITLE OF INVENTION: of Use  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/144,779C  
; FILING DATE: 28-OCT-1993  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/783,517  
; FILING DATE: 28-OCT-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/816,205  
; FILING DATE: 02-JAN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/967,633  
; FILING DATE: 28-OCT-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Storella, John R.  
; REGISTRATION NUMBER: 32,944  
; REFERENCE/DOCKET NUMBER: 15542-5-2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-144-779C-27

Query Match 61.0%; Score 25; DB 3; Length 9;  
Best Local Similarity 44.4%; Pred. No. 2.5e+05;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MLAAQRRV 9  
||| |::|  
Db 1 MLTAZZKAV 9

RESULT 2  
US-09-368-449B-27  
; Sequence 27, Application US/09368449B  
; Patent No. 6346514  
; GENERAL INFORMATION:  
; APPLICANT: Green, Lawrence R.  
; SINACHEVICH, Nicolay V.  
; Ivanov, Vadim T.  
; Mikhailova, Inessa I.  
; Vaskovsky, Boris V.  
; Mikhailov, Alexander N.  
; Khavinson, Vladimir K.  
; Morozov, Vyacheslav G.  
; TITLE OF INVENTION: Immunomodulating Peptides and Methods  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION NUMBER: US/09/368,449B  
; FILING DATE: 04-Aug-1999  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/144,779  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: US 07/816,205  
; FILING DATE: 02-JAN-1992  
; APPLICATION NUMBER: US 07/967,633  
; FILING DATE: 28-OCT-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Storella, John R.  
; REGISTRATION NUMBER: 32,944  
; REFERENCE/DOCKET NUMBER: 15542-5-2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: <Unknown>  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 27:

US-09-368-449B-27  
Query Match 61.0%; Score 25; DB 4; Length 9;

Best Local Similarity 44.4%; Pred. No. 2.5e+05;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 MLAAQRRV 9  
||| |::|  
Db 1 MLTAZZKAV 9

RESULT 3  
US-08-144-779C-24  
; Sequence 24, Application US/08144779C  
; Patent No. 6066622  
; GENERAL INFORMATION:  
; APPLICANT: Green, Lawrence R.  
; SINACHEVICH, Nicolay V.  
; Ivanov, Vadim T.  
; Mikhailova, Inessa I.  
; Vaskovsky, Boris V.  
; Mikhailov, Alexander N.  
; Khavinson, Vladimir K.  
; Morozov, Vyacheslav G.  
; TITLE OF INVENTION: Immunomodulating Peptides and Methods  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION NUMBER: US/08/144,779C  
; FILING DATE: 28-OCT-1993  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/783,517  
; FILING DATE: 28-OCT-1991  
; APPLICATION DATA:  
; APPLICATION NUMBER: US 07/816,205  
; FILING DATE: 02-JAN-1992  
; APPLICATION DATA:  
; APPLICATION NUMBER: US 07/967,633  
; FILING DATE: 28-OCT-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Storella, John R.  
; REGISTRATION NUMBER: 32,944  
; REFERENCE/DOCKET NUMBER: 15542-5-2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 24:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-144-779C-24

Query Match 53.7%; Score 22; DB 3; Length 7;  
Best Local Similarity 42.9%; Pred. No. 2.5e+05;  
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 MLAAQER 7  
||| |::|  
Db 1 MLTAZZK 7

RESULT 4  
US-09-368-449B-24  
; Sequence 24, Application US/09368449B  
; Patent No. 6346514  
; GENERAL INFORMATION:  
; APPLICANT: Green, Lawrence R.  
; Sinackevich, Nicolay V.  
; Ivanov, Vadim T.  
; Mikhalyova, Inessa I.  
; Vaskovsky, Boris V.  
; Mikhaltsov, Alexander N.  
; Khavinson, Vladimir K.  
; Morozov, Vyacheslav G.  
; TITLE OF INVENTION: Immunomodulating Peptides and Methods  
; of Use  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/368,449B  
; FILING DATE: 04-Aug-1999  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/144,779  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: US 07/816,205  
; FILING DATE: 02-JAN-1992  
; APPLICATION NUMBER: US 07/967,633  
; FILING DATE: 28-OCT-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Storella, John R.  
; REGISTRATION NUMBER: 32,944  
; REFERENCE/DOCKET NUMBER: 15542-5-2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 24:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: <Unknown>  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 24:  
US-09-368-449B-24  
Query Match 53.7%; Score 22; DB 4; Length 7;  
Best Local Similarity 42.9%; Pred. No. 2.5e+05;  
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MLLAQER 7  
Db 1 MLTAZK 7  
RESULT 5  
US-08-144-779C-25  
; Sequence 25, Application US/08144779C  
; Patent No. 6066622  
; GENERAL INFORMATION:  
; APPLICANT: Green, Lawrence R.  
; Sinackevich, Nicolay V.

; APPLICANT: Ivanov, Vadim T.  
; APPLICANT: Mikhalyova, Inessa I.  
; APPLICANT: Vaskovsky, Boris V.  
; APPLICANT: Mikhaltsov, Alexander N.  
; APPLICANT: Khavinson, Vladimir K.  
; APPLICANT: Morozov, Vyacheslav G.  
; TITLE OF INVENTION: Immunomodulating Peptides and Methods  
; of Use  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/144,779C  
; FILING DATE: 28-OCT-1993  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/783,517  
; FILING DATE: 28-OCT-1991  
; APPLICATION NUMBER: US 07/816,205  
; FILING DATE: 02-JAN-1992  
; APPLICATION NUMBER: US 07/967,633  
; FILING DATE: 28-OCT-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Storella, John R.  
; REGISTRATION NUMBER: 32,944  
; REFERENCE/DOCKET NUMBER: 15542-5-2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-144-779C-25  
Query Match 53.7%; Score 22; DB 3; Length 8;  
Best Local Similarity 42.9%; Pred. No. 2.5e+05;  
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MLLAQER 7  
Db 1 MLTAZK 7  
RESULT 6  
US-09-368-449B-25  
; Sequence 25, Application US/09368449B  
; Patent No. 6346514  
; GENERAL INFORMATION:  
; APPLICANT: Green, Lawrence R.  
; Sinackevich, Nicolay V.  
; Ivanov, Vadim T.  
; Mikhalyova, Inessa I.  
; Vaskovsky, Boris V.  
; Mikhaltsov, Alexander N.  
; Khavinson, Vladimir K.  
; Morozov, Vyacheslav G.  
; TITLE OF INVENTION: Immunomodulating Peptides and Methods

```
;
;
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/368,449B
; FILING DATE: 04-Aug-1999
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/144,779
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/816,205
; FILING DATE: 02-JAN-1992
; APPLICATION NUMBER: US 07/967,633
; FILING DATE: 28-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Storella, John R.
; REGISTRATION NUMBER: 32,944
; REFERENCE/DOCKET NUMBER: 15542-5-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-368-449B-25

Query Match 53.7%; Score 22; DB 4; Length 8;
Best Local Similarity 42.9%; Pred. No. 2.5e+05;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAAQER 7
Db 1 MLTAZZK 7

RESULT 7
US-08-144-779C-26
; Sequence 26, Application US/08144779C
; Patent No. 6066622
; GENERAL INFORMATION:
; APPLICANT: Green, Lawrence R.
; APPLICANT: Sinackevich, Nicolay V.
; APPLICANT: Ivanov, Vadim T.
; APPLICANT: Mikhailova, Inessa I.
; APPLICANT: Vaskovsky, Boris V.
; APPLICANT: Mikhailov, Alexander N.
; APPLICANT: Khavinson, Vladimir K.
; APPLICANT: Morozov, Vyacheslav G.
; TITLE OF INVENTION: Immunomodulating Peptides and Methods
; TITLE OF INVENTION: of Use
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
```

```
;
;
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/144,779C
; FILING DATE: 28-OCT-1993
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/783,517
; FILING DATE: 28-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/816,205
; FILING DATE: 02-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/967,633
; FILING DATE: 28-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Storella, John R.
; REGISTRATION NUMBER: 32,944
; REFERENCE/DOCKET NUMBER: 15542-5-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-144-779C-26

Query Match 53.7%; Score 22; DB 3; Length 9;
Best Local Similarity 42.9%; Pred. No. 2.5e+05;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAAQER 7
Db 1 MLTAZZK 7

RESULT 8
US-09-368-449B-26
; Sequence 26, Application US/09368449B
; Patent No. 6346514
; GENERAL INFORMATION:
; APPLICANT: Green, Lawrence R.
; APPLICANT: Sinackevich, Nicolay V.
; APPLICANT: Ivanov, Vadim T.
; APPLICANT: Mikhailova, Inessa I.
; APPLICANT: Vaskovsky, Boris V.
; APPLICANT: Mikhailov, Alexander N.
; APPLICANT: Khavinson, Vladimir K.
; APPLICANT: Morozov, Vyacheslav G.
; TITLE OF INVENTION: Immunomodulating Peptides and Methods
; TITLE OF INVENTION: of Use
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

; APPLICATION NUMBER: US/09/368,449B  
; FILING DATE: 04-Aug-1999  
; CLASSIFICATION: <Unknown>  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 08/144,779  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: US 07/816,205  
; FILING DATE: 02-JAN-1992  
; APPLICATION NUMBER: US 07/967,633  
; FILING DATE: 28-OCT-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Storella, John R.  
; REGISTRATION NUMBER: 32,944  
; REFERENCE/DOCKET NUMBER: 15542-5-2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 26:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: <Unknown>  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 26:  
US-09-368-449B-26

Query Match 53.7%; Score 22; DB 4; Length 9;  
Best Local Similarity 42.9%; Pred. No. 2.5e+05;  
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MAAOER 7  
|||  
Db 1 MTAZZK 7

RESULT 9  
US-09-447-399-7  
; Sequence 7, Application US/09447399  
; Patent No. 6525187  
; GENERAL INFORMATION:  
; APPLICANT: El Shaml, Said A.  
; APPLICANT: Menon, Surendra N.  
; APPLICANT: French, Cynthia K.  
; TITLE OF INVENTION: Polynucleotide Encoding Autoantigens Associated With  
; TITLE OF INVENTION: Endometriosis  
; FILE REFERENCE: 107-208US-CIP  
; CURRENT APPLICATION NUMBER: US/09/447,399  
; CURRENT FILING DATE: 1999-11-23  
; EARLIER APPLICATION NUMBER: 60/094,930  
; EARLIER FILING DATE: 1998-07-31  
; EARLIER APPLICATION NUMBER: 09/359,084  
; EARLIER FILING DATE: 1999-07-22  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-447-399-7

Query Match 53.7%; Score 22; DB 4; Length 9;  
Best Local Similarity 66.7%; Pred. No. 2.5e+05;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 AAQERR 8  
:|:|:|  
Db 4 SAEERR 9

RESULT 10  
US-08-055-006-53  
; Sequence 53, Application US/08055006

; Patent No. 6454978  
; GENERAL INFORMATION:  
; APPLICANT: Brostoff, Steven W.  
; Wilson, Darcy B.  
; Smith, Lawrence R.  
; Gold, Daniel P.  
; Carlo, Dennis J.  
; TITLE OF INVENTION: Vaccination and Methods Against Multiple  
; Sclerosis Resulting From Pathogenic Responses By Specific  
; Cell Populations  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CAMPBELL & FLORES LLP  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: United States  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/055,006  
; FILING DATE: 09-Feb-1993  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-IM 9611  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619)535-9001  
; TELEFAX: (619)535-8949  
; INFORMATION FOR SEQ ID NO: 53:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 53:  
US-08-055-006-53

Query Match 51.2%; Score 21; DB 4; Length 6;  
Best Local Similarity 80.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 QRRV 9  
|:|:|  
Db 1 QDRRV 5

RESULT 11  
US-09-627-851B-39  
; Sequence 39, Application US/09627851B  
; Patent No. 6545131  
; GENERAL INFORMATION:  
; APPLICANT: ISSACS, JOHN T.  
; APPLICANT: DENMEADE, SAMUEL R.  
; APPLICANT: CHRISTENSEN, S. BROGGER  
; APPLICANT: LILJA, HANS  
; TITLE OF INVENTION: TISSUE SPECIFIC PRODRUG  
; FILE REFERENCE: 57111 (71699)  
; CURRENT APPLICATION NUMBER: US/09/627,851B  
; CURRENT FILING DATE: 2000-07-28  
; PRIOR APPLICATION NUMBER: 09/588,822  
; PRIOR FILING DATE: 2000-06-07  
; PRIOR APPLICATION NUMBER: 09/081,707  
; PRIOR FILING DATE: 1998-05-19  
; PRIOR APPLICATION NUMBER: 60/047,070  
; PRIOR FILING DATE: 1997-05-19  
; PRIOR APPLICATION NUMBER: 60/080,046  
; PRIOR FILING DATE: 1998-03-30

; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 39  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-627-851B-39

Query Match 51.2%; Score 21; DB 4; Length 7;  
Best Local Similarity 66.7%; Pred. No. 2.5e+05;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 4 AQRRV 9  
Db 2 AQRRRL 7

## RESULT 12

US-09-518-046-78  
; Sequence 78, Application US/09518046  
; Patent No. 6294663  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; TITLE OF INVENTION: Transmembrane Serine Protease Overexpressed  
; FILE REFERENCE: D6192CIP  
; CURRENT APPLICATION NUMBER: US/09/518,046  
; CURRENT FILING DATE: 2000-03-02  
; EARLIER APPLICATION NUMBER: 09/261,416  
; EARLIER FILING DATE: 1999-03-03  
; NUMBER OF SEQ ID NOS: 153  
; SEQ ID NO 78  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; LOCATION: 406...414  
; OTHER INFORMATION: TADG-12 peptide

## US-09-518-046-78

Query Match 51.2%; Score 21; DB 3; Length 9;  
Best Local Similarity 80.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 QRRV 9  
Db 3 QRRRL 7

## RESULT 13

US-09-518-046-95  
; Sequence 95, Application US/09518046  
; Patent No. 6294663  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; TITLE OF INVENTION: Transmembrane Serine Protease Overexpressed  
; FILE REFERENCE: D6192CIP  
; CURRENT APPLICATION NUMBER: US/09/518,046  
; CURRENT FILING DATE: 2000-03-02  
; EARLIER APPLICATION NUMBER: 09/261,416  
; EARLIER FILING DATE: 1999-03-03  
; NUMBER OF SEQ ID NOS: 153  
; SEQ ID NO 95  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; LOCATION: 407...415  
; OTHER INFORMATION: TADG-12 peptide

## US-09-518-046-95

Query Match 51.2%; Score 21; DB 3; Length 9;  
Best Local Similarity 80.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 QRRV 9  
Db 2 QRRRL 6

## RESULT 14

US-09-518-046-146  
; Sequence 146, Application US/09518046  
; Patent No. 6294663  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; TITLE OF INVENTION: Transmembrane Serine Protease Overexpressed  
; FILE REFERENCE: D6192CIP  
; CURRENT APPLICATION NUMBER: US/09/518,046  
; CURRENT FILING DATE: 2000-03-02  
; EARLIER APPLICATION NUMBER: 09/261,416  
; EARLIER FILING DATE: 1999-03-03  
; NUMBER OF SEQ ID NOS: 153  
; SEQ ID NO 146  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; LOCATION: 408...416  
; OTHER INFORMATION: TADG-12 peptide

## US-09-518-046-146

Query Match 51.2%; Score 21; DB 3; Length 9;  
Best Local Similarity 80.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 QRRV 9  
Db 1 QRRRL 5

## RESULT 15

US-09-644-600-59  
; Sequence 59, Application US/09544600  
; Patent No. 6451500  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; TITLE OF INVENTION: Tanimoto, Hirotsu  
; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease  
; FILE REFERENCE: D6064CIP/D  
; CURRENT APPLICATION NUMBER: US/09/644,600  
; CURRENT FILING DATE: 2000-08-23  
; PRIOR APPLICATION NUMBER: 09/421,213  
; PRIOR FILING DATE: 1999-10-20  
; PRIOR APPLICATION NUMBER: 09/027,337  
; PRIOR FILING DATE: 1998-02-20  
; NUMBER OF SEQ ID NOS: 98  
; SEQ ID NO 59  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Residues 686-694 of the TADG-15 protein

## US-09-644-600-59

Query Match 51.2%; Score 21; DB 4; Length 9;  
Best Local Similarity 80.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 QRRV 9

Db       ||||: 9  
          5 QERRL 9

Search completed: October 7, 2003, 14:49:52  
Job time : 21 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:45:41 ; Search time 21 Seconds

(without alignments)  
67.806 Million cell updates/sec

Title: US-09-807-512-26

Perfect score: 41

Sequence: 1 MAAQERRV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues

Total number of hits satisfying chosen parameters: 59020

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA.\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23	56.1	9	11	US-09-799-250-40
2	23	56.1	9	11	US-09-799-250-56
3	23	56.1	9	11	US-09-799-250-121
4	23	56.1	9	11	US-09-799-250-244
5	23	56.1	9	11	US-09-799-250-348
6	23	56.1	9	11	US-09-799-250-560
7	22	53.7	9	15	US-10-172-573-7
8	21	51.2	9	12	US-10-357-175-78
9	21	51.2	9	12	US-10-357-175-95
10	21	51.2	9	12	US-10-357-175-146
11	20	48.8	7	15	US-10-059-261-18
12	20	48.8	7	15	US-10-059-261-126
13	20	48.8	7	15	US-10-353-929-186
14	19	46.3	5	9	US-09-096-749A-63
15	19	46.3	5	11	US-09-903-412-63

16	19	46.3	5	12	US-10-165-155-63	Sequence 63, Appl
17	19	46.3	5	12	US-10-190-162-63	Sequence 63, Appl
18	19	46.3	5	15	US-10-174-717A-63	Sequence 63, Appl
19	19	46.3	9	10	US-09-818-247-15	Sequence 15, Appl
20	19	46.3	9	10	US-09-792-480-17	Sequence 17, Appl
21	19	46.3	9	11	US-09-799-250-31	Sequence 31, Appl
22	19	46.3	9	11	US-09-799-250-158	Sequence 158, App
23	19	46.3	9	11	US-09-799-250-438	Sequence 438, App
24	19	46.3	9	11	US-09-799-250-524	Sequence 524, App
25	19	46.3	9	11	US-09-799-250-620	Sequence 620, App
26	19	46.3	9	12	US-09-932-165-212	Sequence 212, App
27	19	46.3	9	12	US-09-935-384-40	Sequence 40, Appl
28	19	46.3	9	12	US-09-935-384-47	Sequence 47, Appl
29	19	46.3	9	12	US-09-935-384-323	Sequence 323, App
30	19	46.3	9	12	US-09-935-384-432	Sequence 432, App
31	19	46.3	9	12	US-09-935-384-517	Sequence 517, App
32	19	46.3	9	12	US-09-935-384-648	Sequence 648, Appl
33	19	46.3	9	15	US-10-033-662-8	Sequence 8, Appl
34	18	43.9	7	9	US-09-770-940-48	Sequence 48, Appl
35	18	43.9	7	9	US-09-736-076-53	Sequence 53, Appl
36	18	43.9	7	12	US-10-052-578-161	Sequence 161, App
37	18	43.9	7	12	US-10-165-528-48	Sequence 48, Appl
38	18	43.9	7	12	US-10-053-520-161	Sequence 161, App
39	18	43.9	8	9	US-09-736-076-51	Sequence 51, Appl
40	18	43.9	8	9	US-09-736-076-52	Sequence 52, Appl
41	18	43.9	8	14	US-10-007-761-25	Sequence 25, Appl
42	18	43.9	8	14	US-10-007-761-27	Sequence 27, Appl
43	18	43.9	8	14	US-10-007-761-28	Sequence 28, Appl
44	18	43.9	8	14	US-10-007-761-31	Sequence 31, Appl
45	18	43.9	9	9	US-09-796-294-39	Sequence 39, Appl

ALIGNMENTS

RESULT 1  
US-09-799-250-40  
; Sequence 40, Application US/09799250  
; Publication No. US20030032087A1  
; GENERAL INFORMATION:  
; APPLICANT: Pia M. Challita-Eld  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Mary Faris  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: 121P1: A TISSUE SPECIFIC PROTEIN  
; FILE REFERENCE: 129.34US01  
; CURRENT APPLICATION NUMBER: US/09/799,250  
; CURRENT FILING DATE: 2001-02-08  
; NUMBER OF SEQ ID NOS: 719  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 40  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-799-250-40

Query Match 56.1%; Score 23; DB 11; Length 9;  
Best Local Similarity 57.1%; Pred. No. 5.2e+05;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAAQERR 8  
Db 1 LSAEEKR 7

RESULT 2  
US-09-799-250-56  
; Sequence 56, Application US/09799250  
; Publication No. US20030032087A1

GENERAL INFORMATION:  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Mary Paris  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Ava Jakobovits  
; TITLE OF INVENTION: 121P1F1: A TISSUE SPECIFIC PROTEIN  
; FILE REFERENCE: 129.34US01  
; CURRENT APPLICATION NUMBER: US/09/799,250  
; CURRENT FILING DATE: 2001-02-08  
; NUMBER OF SEQ ID NOS: 719  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 56  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-799-250-56

Query Match 56.1%; Score 23; DB 11; Length 9;  
Best Local Similarity 57.1%; Pred. No. 5.2e+05;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAAQERR 8  
|:|:|:  
Db 3 LSAEKR 9

RESULT 3  
US-09-799-250-121  
; Sequence 121, Application US/09799250  
; Publication No. US20030032087A1  
; GENERAL INFORMATION:  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Mary Paris  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Ava Jakobovits  
; TITLE OF INVENTION: 121P1F1: A TISSUE SPECIFIC PROTEIN  
; FILE REFERENCE: 129.34US01  
; CURRENT APPLICATION NUMBER: US/09/799,250  
; CURRENT FILING DATE: 2001-02-08  
; NUMBER OF SEQ ID NOS: 719  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 121  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-799-250-121

Query Match 56.1%; Score 23; DB 11; Length 9;  
Best Local Similarity 57.1%; Pred. No. 5.2e+05;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAAQERR 8  
|:|:|:  
Db 2 LSAEKR 8

RESULT 4  
US-09-799-250-244  
; Sequence 244, Application US/09799250  
; Publication No. US20030032087A1  
; GENERAL INFORMATION:  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Arthur B. Raitano

; APPLICANT: Mary Paris  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Ava Jakobovits  
; TITLE OF INVENTION: 121P1F1: A TISSUE SPECIFIC PROTEIN  
; FILE REFERENCE: 129.34US01  
; CURRENT APPLICATION NUMBER: US/09/799,250  
; CURRENT FILING DATE: 2001-02-08  
; NUMBER OF SEQ ID NOS: 719  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 244  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-799-250-244

Query Match 56.1%; Score 23; DB 11; Length 9;  
Best Local Similarity 57.1%; Pred. No. 5.2e+05;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAAQERR 8  
|:|:|:  
Db 2 LSAEKR 8

RESULT 5  
US-09-799-250-348  
; Sequence 348, Application US/09799250  
; Publication No. US20030032087A1  
; GENERAL INFORMATION:  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Mary Paris  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Ava Jakobovits  
; TITLE OF INVENTION: 121P1F1: A TISSUE SPECIFIC PROTEIN  
; FILE REFERENCE: 129.34US01  
; CURRENT APPLICATION NUMBER: US/09/799,250  
; CURRENT FILING DATE: 2001-02-08  
; NUMBER OF SEQ ID NOS: 719  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 348  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-799-250-348

Query Match 56.1%; Score 23; DB 11; Length 9;  
Best Local Similarity 57.1%; Pred. No. 5.2e+05;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAAQERR 8  
|:|:|:  
Db 3 LSAEKR 9

RESULT 6  
US-09-799-250-560  
; Sequence 560, Application US/09799250  
; Publication No. US20030032087A1  
; GENERAL INFORMATION:  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Mary Paris  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Ava Jakobovits  
; TITLE OF INVENTION: 121P1F1: A TISSUE SPECIFIC PROTEIN  
; FILE REFERENCE: 129.34US01  
; CURRENT APPLICATION NUMBER: US/09/799,250  
; CURRENT FILING DATE: 2001-02-08  
; NUMBER OF SEQ ID NOS: 719  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 560  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-799-250-560

Query Match 56.1%; Score 23; DB 11; Length 9;  
Best Local Similarity 57.1%; Pred. No. 5.2e+05;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAAQERR 8  
|:|:|:  
Db 3 LSAEKR 9

; FILE REFERENCE: 129.34US01  
; CURRENT APPLICATION NUMBER: US/09/799,250  
; CURRENT FILING DATE: 2001-02-08  
; NUMBER OF SEQ ID NOS: 719  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 560  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-799-250-560

Query Match 56.1%; Score 23; DB 11; Length 9;  
Best Local Similarity 57.1%; Pred. No. 5.2e+05;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAAQERR 8  
|:|:|:|  
Db 2 LSAEERR 8

RESULT 7  
US-10-172-573-7  
; Sequence 7, Application US/10172573  
; Publication No. US20030083253A1  
; GENERAL INFORMATION:  
; APPLICANT: El Shami, Said A.  
; APPLICANT: Menon, Surendra N.  
; APPLICANT: French, Cynthia K.  
; TITLE OF INVENTION: Polynucleotide Encoding Autoantigens Associated With  
; FILE REFERENCE: Endometriosis  
; CURRENT APPLICATION NUMBER: US/10/172,573  
; CURRENT FILING DATE: 2002-06-13  
; PRIOR APPLICATION NUMBER: US/09/447,399  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: 60/094,930  
; PRIOR FILING DATE: 1998-07-31  
; PRIOR APPLICATION NUMBER: 09/359,084  
; PRIOR FILING DATE: 1999-07-22  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-172-573-7

Query Match 53.7%; Score 22; DB 15; Length 9;  
Best Local Similarity 66.7%; Pred. No. 5.2e+05;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAQERR 8  
|:|:|:|  
Db 4 SAEERR 9

RESULT 8  
US-10-357-175-78  
; Sequence 78, Application US/10357175  
; Publication No. US20030170707A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; TITLE OF INVENTION: Transmembrane Serine Protease Overexpressed  
; FILE REFERENCE: D6192CIP/D/CIP  
; CURRENT APPLICATION NUMBER: US/10/357,175  
; CURRENT FILING DATE: 2003-02-03  
; PRIOR APPLICATION NUMBER: 09/650,371  
; PRIOR FILING DATE: 2000-08-28  
; NUMBER OF SEQ ID NOS: 158  
; SEQ ID NO 78  
; LENGTH: 9  
; TYPE: PRT

; ORGANISM: Homo sapiens  
; FEATURE:  
; LOCATION: 406...414  
; OTHER INFORMATION: TAGD-12 peptide  
US-10-357-175-78

Query Match 51.2%; Score 21; DB 12; Length 9;  
Best Local Similarity 80.0%; Pred. No. 5.2e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 QERRV 9  
|:|:|:|  
Db 3 QERRL 7

RESULT 9  
US-10-357-175-95  
; Sequence 95, Application US/10357175  
; Publication No. US20030170707A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; TITLE OF INVENTION: Transmembrane Serine Protease Overexpressed  
; FILE REFERENCE: D6192CIP/D/CIP  
; CURRENT APPLICATION NUMBER: US/10/357,175  
; CURRENT FILING DATE: 2003-02-03  
; PRIOR APPLICATION NUMBER: 09/650,371  
; PRIOR FILING DATE: 2000-08-28  
; NUMBER OF SEQ ID NOS: 158  
; SEQ ID NO 95  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; LOCATION: 407...415  
; OTHER INFORMATION: TAGD-12 peptide  
US-10-357-175-95

Query Match 51.2%; Score 21; DB 12; Length 9;  
Best Local Similarity 80.0%; Pred. No. 5.2e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 QERRV 9  
|:|:|:|  
Db 2 QERRL 6

RESULT 10  
US-10-357-175-146  
; Sequence 146, Application US/10357175  
; Publication No. US20030170707A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; TITLE OF INVENTION: Transmembrane Serine Protease Overexpressed  
; FILE REFERENCE: D6192CIP/D/CIP  
; CURRENT APPLICATION NUMBER: US/10/357,175  
; CURRENT FILING DATE: 2003-02-03  
; PRIOR APPLICATION NUMBER: 09/650,371  
; PRIOR FILING DATE: 2000-08-28  
; NUMBER OF SEQ ID NOS: 158  
; SEQ ID NO 146  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; LOCATION: 408...416  
; OTHER INFORMATION: TAGD-12 peptide  
US-10-357-175-146

Query Match 51.2%; Score 21; DB 12; Length 9;  
Best Local Similarity 80.0%; Pred. No. 5.2e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 QERRV 9  
!!!!:  
Db 1 QERRL 5

## RESULT 11

US-10-059-261-18  
; Sequence 18, Application US/10059261  
; Publication No. US20030077826A1  
; GENERAL INFORMATION:  
; APPLICANT: EDELMAN, LENA  
; APPLICANT: JACOTOT, ETIENNE DANIEL FRANCOIS  
; APPLICANT: BRIAND, JEAN-PAUL  
; TITLE OF INVENTION: CHIMERIC MOLECULES CONTAINING A MODULE ABLE TO TARGET  
; TITLE OF INVENTION: SPECIFIC CELLS AND A MODULE REGULATING THE APOPTOTIC  
; TITLE OF INVENTION: FUNCTION OF THE PERMEABILITY TRANSITION PORE COMPLEX  
; TITLE OF INVENTION: (PTPC)  
; FILE REFERENCE: 03495.0216  
; CURRENT APPLICATION NUMBER: US/10/059.261  
; CURRENT FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/265,594  
; PRIOR FILING DATE: 2001-02-02  
; NUMBER OF SEQ ID NOS: 325  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: vector pACgp67-ScFv461  
US-10-059-261-18

Query Match 48.8%; Score 20; DB 15; Length 7;  
Best Local Similarity 42.9%; Pred. No. 5.2e+05;  
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 AAQERRV 9  
!!!!:  
Db 1 SAESRRI 7

## RESULT 12

US-10-059-261-126  
; Sequence 126, Application US/10059261  
; Publication No. US20030077826A1  
; GENERAL INFORMATION:  
; APPLICANT: EDELMAN, LENA  
; APPLICANT: JACOTOT, ETIENNE DANIEL FRANCOIS  
; APPLICANT: BRIAND, JEAN-PAUL  
; TITLE OF INVENTION: CHIMERIC MOLECULES CONTAINING A MODULE ABLE TO TARGET  
; TITLE OF INVENTION: SPECIFIC CELLS AND A MODULE REGULATING THE APOPTOTIC  
; TITLE OF INVENTION: FUNCTION OF THE PERMEABILITY TRANSITION PORE COMPLEX  
; TITLE OF INVENTION: (PTPC)  
; FILE REFERENCE: 03495.0216  
; CURRENT APPLICATION NUMBER: US/10/059.261  
; CURRENT FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: 60/265,594  
; PRIOR FILING DATE: 2001-02-02  
; NUMBER OF SEQ ID NOS: 325  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 126  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: vector pACgp67-ScFv350  
US-10-059-261-126

Query Match 48.8%; Score 20; DB 15; Length 7;  
Best Local Similarity 42.9%; Pred. No. 5.2e+05;

Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
Qy 3 AAQERRV 9  
!!!!:  
Db 1 SAESRRI 7

## RESULT 13

US-10-353-929-186  
; Sequence 186, Application US/10353929  
; Publication No. US20030175288A1  
; GENERAL INFORMATION:  
; APPLICANT: ITOH, KYOGO  
; TITLE OF INVENTION: Tumor antigen  
; FILE REFERENCE: GP01-1024  
; CURRENT APPLICATION NUMBER: US/10/353.929  
; CURRENT FILING DATE: 2003-01-30  
; PRIOR APPLICATION NUMBER: JP P2000-231814  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 197  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 186  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Designed peptide based on the amino acid sequence of SEQ ID NO  
US-10-353-929-186

Query Match 48.8%; Score 20; DB 12; Length 9;  
Best Local Similarity 44.4%; Pred. No. 5.2e+05;  
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MLAAQERRV 9  
!!!!:  
Db 1 LLVASQRSV 9

## RESULT 14

US-09-096-749A-63  
; Sequence 63, Application US/09096749A  
; Patent No. US20020019517A1  
; GENERAL INFORMATION:  
; APPLICANT: Koieda, Shohel  
; TITLE OF INVENTION: ARTIFICIAL ANTIBODY POLYPEPTIDES  
; NUMBER OF SEQUENCES: 118  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth P.A.  
; STREET: 121 South Eighth Street, Ste. 1600  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0b  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/096,749A  
; FILING DATE: June 12, 1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ann S. Viksna  
; REGISTRATION NUMBER: 37,748  
; REFERENCE/DOCKET NUMBER: 109.034US1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (612) 373-6900  
; TELEFAX: (612) 339-3061  
; INFORMATION FOR SEQ ID NO: 63:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 5 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
; ORIGINAL SOURCE:  
US-09-096-749A-63

Query Match 46.3%; Score 19; DB 9; Length 5;  
Best Local Similarity 100.0%; Pred. No. 5.2e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ERRV 9  
|||||  
Db 1 ERRV 4

## RESULT 15

US-09-903-412-63  
; Sequence 63, Application US/09903412  
; Publication No. US20030027319A1  
; GENERAL INFORMATION:  
; APPLICANT: Koide, Shohei  
; TITLE OF INVENTION: ARTIFICIAL ANTIBODY POLYPEPTIDES  
; FILE REFERENCE: 109, 050US1  
; CURRENT APPLICATION NUMBER: US/09/903,412  
; CURRENT FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: US 60/217,474  
; PRIOR FILING DATE: 2000-07-11  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 63  
; LENGTH: 5  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: The sequence of the BC loop of clone pLB24.8.  
US-09-903-412-63

Query Match 46.3%; Score 19; DB 11; Length 5;  
Best Local Similarity 100.0%; Pred. No. 5.2e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ERRV 9  
|||||  
Db 1 ERRV 4

Search completed: October 7, 2003, 14:50:41  
Job time : 22 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 14:46:12 ; Search time 277.5 seconds  
(without alignments)  
29.511 Million cell updates/sec

Title: US-09-807-512-26  
Perfect score: 41  
Sequence: 1 MAAQERRV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues

Total number of hits satisfying chosen parameters: 435500

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main:\*

1: /cgn2\_6/ptodata/1/paa/PCTUS\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep.\*  
8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep.\*  
9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep.\*  
10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep.\*  
11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep.\*  
12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep.\*  
13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep.\*  
14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep.\*  
15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep.\*  
16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep.\*  
17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep.\*  
18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep.\*  
19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep.\*  
20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep.\*  
21: /cgn2\_6/ptodata/1/paa/US097A\_COMB.pep.\*  
22: /cgn2\_6/ptodata/1/paa/US097B\_COMB.pep.\*  
23: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep.\*  
24: /cgn2\_6/ptodata/1/paa/US099A\_COMB.pep.\*  
25: /cgn2\_6/ptodata/1/paa/US099B\_COMB.pep.\*  
26: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep.\*  
27: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep.\*  
28: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep.\*  
29: /cgn2\_6/ptodata/1/paa/US103\_COMB.pep.\*  
30: /cgn2\_6/ptodata/1/paa/US104\_COMB.pep.\*  
31: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep.\*  
32: /cgn2\_6/ptodata/1/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	ID	Description
1	41	100.0	9 23	US-09-807-512-26 Sequence 26, Appl

2	32	78.0	9	19	US-09-529-206-46	Sequence 46, Appl
3	32	78.0	9	19	US-09-529-206A-46	Sequence 46, Appl
4	32	78.0	9	19	US-09-529-206A-123	Sequence 123, App
5	32	78.0	9	19	US-09-529-206B-46	Sequence 46, Appl
6	32	78.0	9	19	US-09-529-206B-123	Sequence 123, App
7	32	78.0	9	19	US-09-529-206B-46	Sequence 46, Appl
8	32	78.0	9	19	US-09-529-206D-123	Sequence 123, App
9	25	61.0	9	5	US-08-144-779-27	Sequence 27, Appl
10	24	58.5	9	27	US-10-119-536A-181	Sequence 181, App
11	23	56.1	9	22	US-09-799-250-40	Sequence 40, Appl
12	23	56.1	9	22	US-09-799-250-56	Sequence 56, Appl
13	23	56.1	9	22	US-09-799-250-121	Sequence 121, App
14	23	56.1	9	22	US-09-799-250-244	Sequence 244, App
15	23	56.1	9	22	US-09-799-250-348	Sequence 348, App
16	23	56.1	9	22	US-09-799-250-560	Sequence 560, App
17	23	56.1	9	22	US-09-799-250A-40	Sequence 40, Appl
18	23	56.1	9	22	US-09-799-250A-56	Sequence 56, Appl
19	23	56.1	9	22	US-09-799-250A-121	Sequence 121, App
20	23	56.1	9	22	US-09-799-250A-244	Sequence 244, App
21	23	56.1	9	22	US-09-799-250A-348	Sequence 348, App
22	23	56.1	9	22	US-09-799-250A-560	Sequence 560, App
23	22	53.7	7	5	US-08-144-779-24	Sequence 24, Appl
24	22	53.7	8	5	US-08-144-779-25	Sequence 25, Appl
25	22	53.7	8	7	US-08-370-838-16	Sequence 16, Appl
26	22	53.7	8	10	US-08-686-004-16	Sequence 16, Appl
27	22	53.7	9	5	US-08-144-779-26	Sequence 26, Appl
28	22	53.7	9	27	US-10-172-573-7	Sequence 7, Appl
29	21	51.2	6	13	US-08-984-098-229	Sequence 229, App
30	21	51.2	7	1	PCT-US00-40496-43	Sequence 43, Appl
31	21	51.2	7	20	US-09-627-600B-43	Sequence 43, Appl
32	21	51.2	8	1	PCT-US00-40496-93	Sequence 93, Appl
33	21	51.2	8	4	US-08-045-394A-32	Sequence 32, Appl
34	21	51.2	8	20	US-09-627-600B-93	Sequence 93, Appl
35	21	51.2	9	1	PCT-US00-05612-78	Sequence 78, Appl
36	21	51.2	9	1	PCT-US00-05612-95	Sequence 95, Appl
37	21	51.2	9	1	PCT-US00-05612-146	Sequence 146, App
38	21	51.2	9	1	PCT-US00-29095-59	Sequence 59, Appl
39	21	51.2	9	18	US-09-421-213-59	Sequence 59, Appl
40	21	51.2	9	20	US-09-650-371-78	Sequence 78, Appl
41	21	51.2	9	20	US-09-650-371-95	Sequence 95, Appl
42	21	51.2	9	20	US-09-650-371-146	Sequence 146, App
43	21	51.2	9	20	US-09-654-600A-59	Sequence 59, Appl
44	21	51.2	9	29	US-10-357-175-78	Sequence 78, Appl
45	21	51.2	9	29	US-10-357-175-95	Sequence 95, Appl

ALIGNMENTS

RESULT 1

US-09-807-512-26  
; Sequence 26, Application US/09807512  
; GENERAL INFORMATION:  
; APPLICANT: Schrier, Peter I.  
; APPLICANT: Aarnoudse, Corlien  
; APPLICANT: Heider, Karl-Heinz  
; APPLICANT: Klade, Christoph  
; TITLE OF INVENTION: Camel, An Alternative Translation Product of the Tumor Antigen-Lage 1  
; FILE REFERENCE: 0652.220000  
; CURRENT APPLICATION NUMBER: US/09/807,512  
; CURRENT FILING DATE: 2001-04-16  
; PRIOR APPLICATION NUMBER: PCT/EP99/07832  
; PRIOR FILING DATE: 1999-10-15  
; PRIOR APPLICATION NUMBER: EP 98119583.7  
; PRIOR FILING DATE: 1998-10-16  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 26  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-807-512-26

```
Query Match      100.0%; Score 41; DB 23; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.3e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAAQRRV 9
    |||||
Db 1 MLAAQRRV 9

RESULT 2
US-09-529-206-46
; Sequence 46, Application US/09529206
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America as
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; FILE REFERENCE: 20264269US1
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US/09/529,206
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206-46

Query Match      78.0%; Score 32; DB 19; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.3e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 AAQRRV 9
    |||||
Db 1 AAQRRV 7

RESULT 3
US-09-529-206A-46
; Sequence 46, Application US/09529206A
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; FILE REFERENCE: 20264269US1
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US/09/529,206A
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: PCT/US98/19609
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206A-46

Query Match      78.0%; Score 32; DB 19; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.3e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 AAQRRV 9
    |||||
Db 1 AAQRRV 7

RESULT 4
US-09-529-206A-123
; Sequence 123, Application US/09529206A
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; FILE REFERENCE: 20264269US1
; CURRENT FILING DATE: 2000-06-13
```

```
; Sequence 123, Application US/09529206A
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; FILE REFERENCE: 20264269US1
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: PCT/US98/19609
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 123
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206A-123

Query Match      78.0%; Score 32; DB 19; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.3e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAAQRR 7
    |||||
Db 3 MLAAQRR 9

RESULT 5
US-09-529-206B-46
; Sequence 46, Application US/09529206B
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; FILE REFERENCE: 20264269US1
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: PCT/US98/19609
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US60/061,428
; PRIOR FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-206B-46

Query Match      78.0%; Score 32; DB 19; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.3e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 AAQRRV 9
    |||||
Db 1 AAQRRV 7

RESULT 6
US-09-529-206B-123
; Sequence 123, Application US/09529206B
; GENERAL INFORMATION:
; APPLICANT: Wang Rong, Fu
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene
; FILE REFERENCE: 20264269US1
; CURRENT FILING DATE: 2000-06-13
```



; PRIOR APPLICATION NUMBER: PCT/US98/19609  
; PRIOR FILING DATE: 1998-09-21  
; PRIOR APPLICATION NUMBER: US60/061.428  
; PRIOR FILING DATE: 1997-10-08  
; NUMBER OF SEQ ID NOS: 127  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 123  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-206B-123

Query Match 78.0%; Score 32; DB 19; Length 9;  
Best Local Similarity 100.0%; Pred. No. 5.3e+06;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAQER 7  
Db 3 MAAQER 9

RESULT 7  
US-09-529-206D-46  
; SEQUENCE 46, Application US/09529206D  
; GENERAL INFORMATION:  
; APPLICANT: Wang Rong, Fu  
; APPLICANT: Rosenberg, Steven  
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene  
; TITLE OF INVENTION: Encoding Same  
; FILE REFERENCE: 218791  
; CURRENT APPLICATION NUMBER: US/09/529,206D  
; CURRENT FILING DATE: 2000-06-13  
; PRIOR APPLICATION NUMBER: PCT/US98/19609  
; PRIOR FILING DATE: 1998-09-21  
; PRIOR APPLICATION NUMBER: US60/061.428  
; PRIOR FILING DATE: 1997-10-08  
; NUMBER OF SEQ ID NOS: 127  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 46  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-206D-46

Query Match 78.0%; Score 32; DB 19; Length 9;  
Best Local Similarity 100.0%; Pred. No. 5.3e+06;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAQERRV 9  
Db 1 AAQERRV 7

RESULT 8  
US-09-529-206D-123  
; SEQUENCE 123, Application US/09529206D  
; GENERAL INFORMATION:  
; APPLICANT: Wang Rong, Fu  
; APPLICANT: Rosenberg, Steven  
; TITLE OF INVENTION: Novel Human Cancer Antigen NY ESO-1/CAG-3 and Gene  
; TITLE OF INVENTION: Encoding Same  
; FILE REFERENCE: 218791  
; CURRENT APPLICATION NUMBER: US/09/529,206D  
; CURRENT FILING DATE: 2000-06-13  
; PRIOR APPLICATION NUMBER: PCT/US98/19609  
; PRIOR FILING DATE: 1998-09-21  
; PRIOR APPLICATION NUMBER: US60/061.428  
; PRIOR FILING DATE: 1997-10-08  
; NUMBER OF SEQ ID NOS: 127  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 123  
; LENGTH: 9  
; TYPE: PRT

; ORGANISM: Homo sapiens  
US-09-529-206D-123

Query Match 78.0%; Score 32; DB 19; Length 9;  
Best Local Similarity 100.0%; Pred. No. 5.3e+06;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAQER 7  
Db 3 MAAQER 9

RESULT 9  
US-08-144-779-27  
; SEQUENCE 27, Application US/08144779  
; GENERAL INFORMATION:  
; APPLICANT: Green, Lawrence R.  
; APPLICANT: Sinackevich, Nicolay V.  
; APPLICANT: Ivanov, Vadim T.  
; APPLICANT: Mikhailova, Inessa I.  
; APPLICANT: Vaskovsky, Boris V.  
; APPLICANT: Mikhailsov, Alexander N.  
; APPLICANT: Khavinson, Vladimir K.  
; APPLICANT: Morozov, Vyacheslav G.  
; TITLE OF INVENTION: Pharmaceutical Lysine-Containing  
; TITLE OF INVENTION: Polypeptide Compositions and Methods of Use Thereof  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/144,779  
; FILING DATE: 28-OCT-1993  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/783,517  
; FILING DATE: 28-OCT-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/816,205  
; FILING DATE: 02-JAN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/967,633  
; FILING DATE: 28-OCT-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Storella, John R.  
; REGISTRATION NUMBER: 32,944  
; REFERENCE/DOCKET NUMBER: 15542-5-2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-144-779-27

Query Match 61.0%; Score 25; DB 5; Length 9;  
Best Local Similarity 44.4%; Pred. No. 5.3e+06;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAAQERRV 9

```
Db      1 MLTAZZKAV 9
|||:::|
RESULT 10
US-10-119-536A-181
; Sequence 181, Application US/10119536A
; GENERAL INFORMATION:
; APPLICANT: Shaw, Allan Christian
; APPLICANT: Vandahl, Brian Berg
; TITLE OF INVENTION: Method for Identification of Proteins from Intracellular Bacteria
; FILE REFERENCE: 4305/LJ09IUSI
; CURRENT APPLICATION NUMBER: US/10/119,536A
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: DK PA 200100581
; PRIOR FILING DATE: 2001-04-09
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 181
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
; NAME/KEY: BINDING
; LOCATION: (1)...(9)
; OTHER INFORMATION: CT841 immunogenic peptide
US-10-119-536A-181

Query Match      58.5%; Score 24; DB 27; Length 9;
Best Local Similarity 75.0%; Pred. No. 5.3e+06;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 LAAQERRV 9
Db      2 LAALERNV 9
||| |||

RESULT 11
US-09-799-250-40
; Sequence 40, Application US/09799250
; GENERAL INFORMATION:
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Rene S. Hubert
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Arthur B. Raitano
; APPLICANT: Mary Faris
; APPLICANT: Daniel E.H. Afar
; TITLE OF INVENTION: 121P1f1: A TISSUE SPECIFIC PROTEIN
; FILE REFERENCE: 129.34US01
; CURRENT APPLICATION NUMBER: US/09/799,250
; CURRENT FILING DATE: 2001-02-08
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-799-250-40

Query Match      56.1%; Score 23; DB 22; Length 9;
Best Local Similarity 57.1%; Pred. No. 5.3e+06;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      2 LAAQERR 8
Db      1 LSAEKR 7
|||:::|

RESULT 12
US-09-799-250-56
; Sequence 56, Application US/09799250
; GENERAL INFORMATION:
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Rene S. Hubert
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Arthur B. Raitano
; APPLICANT: Mary Faris
; APPLICANT: Daniel E.H. Afar
; TITLE OF INVENTION: 121P1f1: A TISSUE SPECIFIC PROTEIN
; FILE REFERENCE: 129.34US01
; CURRENT APPLICATION NUMBER: US/09/799,250
; CURRENT FILING DATE: 2001-02-08
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 56
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-799-250-56

Query Match      56.1%; Score 23; DB 22; Length 9;
Best Local Similarity 57.1%; Pred. No. 5.3e+06;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      2 LAAQERR 8
Db      2 LSAEKR 8
|||:::|

RESULT 13
US-09-799-250-121
; Sequence 121, Application US/09799250
; GENERAL INFORMATION:
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Rene S. Hubert
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Arthur B. Raitano
; APPLICANT: Mary Faris
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 121P1f1: A TISSUE SPECIFIC PROTEIN
; FILE REFERENCE: 129.34US01
; CURRENT APPLICATION NUMBER: US/09/799,250
; CURRENT FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 719
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 121
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-799-250-121

Query Match      56.1%; Score 23; DB 22; Length 9;
Best Local Similarity 57.1%; Pred. No. 5.3e+06;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      2 LAAQERR 8
Db      3 LSAEKR 9
|||:::|

RESULT 14
US-09-799-250-244
; Sequence 244, Application US/09799250
; GENERAL INFORMATION:
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Rene S. Hubert
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Arthur B. Raitano
; APPLICANT: Mary Faris
; APPLICANT: Daniel E.H. Afar
; TITLE OF INVENTION: 121P1f1: A TISSUE SPECIFIC PROTEIN
; FILE REFERENCE: 129.34US01
; CURRENT APPLICATION NUMBER: US/09/799,250
; CURRENT FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 719
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 244
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-799-250-244
```

; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: 121PIF1: A TISSUE SPECIFIC PROTEIN  
; FILE REFERENCE: 129.34US01  
; CURRENT APPLICATION NUMBER: US/09/799,250  
; CURRENT FILING DATE: 2001-02-08  
; NUMBER OF SEQ ID NOS: 719  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 244  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-799-250-244

Query Match 56.1%; Score 23; DB 22; Length 9;  
Best Local Similarity 57.1%; Pred. No. 5.3e+06;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
QY 2 LAAQERR 8  
DB 2 LSAAEKR 8

RESULT 15  
US-09-799-250-348  
; Sequence 348, Application US/09799250  
; GENERAL INFORMATION:  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Mary Faris  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: 121PIF1: A TISSUE SPECIFIC PROTEIN  
; FILE REFERENCE: 129.34US01  
; CURRENT APPLICATION NUMBER: US/09/799,250  
; CURRENT FILING DATE: 2001-02-08  
; NUMBER OF SEQ ID NOS: 719  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 348  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-799-250-348

Query Match 56.1%; Score 23; DB 22; Length 9;  
Best Local Similarity 57.1%; Pred. No. 5.3e+06;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
QY 2 LAAQERR 8  
DB 3 LSAAEKR 9

Search completed: October 7, 2003, 15:00:04  
Job time : 278.5 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model.

Run on: October 7, 2003, 14:48:22 ; Search time 14.5 Seconds  
(without alignments)  
18.078 Million cell updates/sec

Title: US-09-807-512-26  
Perfect score: 41  
Sequence: 1 LMAQERRV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 136899 seqs, 29126274 residues

Total number of hits satisfying chosen parameters: 17417

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23	56.1	9	5	US-09-799-250B-40
2	23	56.1	9	5	US-09-799-250B-56
3	23	56.1	9	5	US-09-799-250B-121
4	23	56.1	9	5	US-09-799-250B-244
5	23	56.1	9	5	US-09-799-250B-348
6	23	56.1	9	5	US-09-799-250B-560
7	21	51.2	7	5	US-09-627-600C-43
8	21	51.2	7	5	US-09-627-600D-43
9	21	51.2	8	5	US-09-627-600C-93
10	21	51.2	8	5	US-09-627-600C-124
11	20	48.8	8	5	US-09-627-600D-124
12	20	48.8	8	5	US-09-627-600D-124
13	20	48.8	8	5	US-09-390-061D-251
14	20	48.8	8	5	US-09-390-061D-2053
15	20	48.8	8	5	US-09-390-061D-2240
16	20	48.8	8	5	US-09-390-061D-2256
17	20	48.8	8	5	US-09-390-061D-2883
18	20	48.8	8	5	US-09-390-061D-2898
19	20	48.8	8	5	US-09-627-600C-103
20	20	48.8	8	5	US-09-627-600D-103
21	20	48.8	9	1	PCT-US02-27760-11
22	20	48.8	9	1	PCT-US02-27760-32
23	20	48.8	9	5	US-09-390-061D-616
24	20	48.8	9	5	US-09-390-061D-698
25	20	48.8	9	5	US-09-390-061D-2054
26	20	48.8	9	5	US-09-390-061D-2223

ALIGNMENTS

RESULT 1  
US-09-799-250B-40  
; Sequence 40, Application US/09799250B  
; GENERAL INFORMATION:  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Mary Faris  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: 121P1F1: A TISSUE SPECIFIC PROTEIN  
; FILE REFERENCE: 129.34US01 (511582003400)  
; CURRENT APPLICATION NUMBER: US/09/799,250B  
; CURRENT FILING DATE: 2003-07-14  
; NUMBER OF SEQ ID NOS: 721  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 40  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-799-250B-40

Query Match 56.1%; Score 23; DB 5; Length 9;  
Best Local Similarity 57.1%; Pred. No. 1.2e+05;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAAQERR 8  
|:|:|:|  
Db 1 LSAEKR 7

RESULT 2  
US-09-799-250B-56  
; Sequence 56, Application US/09799250B  
; GENERAL INFORMATION:  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Mary Faris  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: 121P1F1: A TISSUE SPECIFIC PROTEIN  
; FILE REFERENCE: 129.34US01 (511582003400)  
; CURRENT APPLICATION NUMBER: US/09/799,250B  
; CURRENT FILING DATE: 2003-07-14

Sequence 2241, Ap  
Sequence 2257, Ap  
Sequence 2700, Ap  
Sequence 2871, Ap  
Sequence 3515, Ap  
Sequence 31, Appl  
Sequence 158, App  
Sequence 438, App  
Sequence 524, App  
Sequence 620, App  
Sequence 78, Appl  
Sequence 78, Appl  
Sequence 249, App  
Sequence 335, App  
Sequence 563, App  
Sequence 32, Appl  
Sequence 61, Appl  
Sequence 171, App  
Sequence 39, Appl

```
; NUMBER OF SEQ ID NOS: 721
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 56
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-799-250B-56

Query Match          56.1%; Score 23; DB 5; Length 9;
Best Local Similarity 57.1%; Pred. No. 1.2e+05;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAAQERR 8
Db 3 LSAEKR 9

RESULT 3
US-09-799-250B-121
; Sequence 121, Application US/09799250B
; GENERAL INFORMATION:
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Rene S. Hubert
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Arthur B. Raitano
; APPLICANT: Mary Faris
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 121P1: A TISSUE SPECIFIC PROTEIN
; FILE REFERENCE: 129.34US01 (511582003400)
; CURRENT APPLICATION NUMBER: US/09799,250B
; CURRENT FILING DATE: 2003-07-14
; NUMBER OF SEQ ID NOS: 721
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 121
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-799-250B-121

Query Match          56.1%; Score 23; DB 5; Length 9;
Best Local Similarity 57.1%; Pred. No. 1.2e+05;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAAQERR 8
Db 2 LSAEKR 8

RESULT 4
US-09-799-250B-244
; Sequence 244, Application US/09799250B
; GENERAL INFORMATION:
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Rene S. Hubert
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Arthur B. Raitano
; APPLICANT: Mary Faris
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 121P1: A TISSUE SPECIFIC PROTEIN
; FILE REFERENCE: 129.34US01 (511582003400)
; CURRENT APPLICATION NUMBER: US/09799,250B
; CURRENT FILING DATE: 2003-07-14
; NUMBER OF SEQ ID NOS: 721
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 244
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-799-250B-244

Query Match          56.1%; Score 23; DB 5; Length 9;
Best Local Similarity 57.1%; Pred. No. 1.2e+05;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAAQERR 8
Db 2 LSAEKR 8

RESULT 5
US-09-799-250B-348
; Sequence 348, Application US/09799250B
; GENERAL INFORMATION:
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Rene S. Hubert
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Arthur B. Raitano
; APPLICANT: Mary Faris
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 121P1: A TISSUE SPECIFIC PROTEIN
; FILE REFERENCE: 129.34US01 (511582003400)
; CURRENT APPLICATION NUMBER: US/09799,250B
; CURRENT FILING DATE: 2003-07-14
; NUMBER OF SEQ ID NOS: 721
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 348
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-799-250B-348

Query Match          56.1%; Score 23; DB 5; Length 9;
Best Local Similarity 57.1%; Pred. No. 1.2e+05;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAAQERR 8
Db 3 LSAEKR 9

RESULT 6
US-09-799-250B-560
; Sequence 560, Application US/09799250B
; GENERAL INFORMATION:
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Rene S. Hubert
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Arthur B. Raitano
; APPLICANT: Mary Faris
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 121P1: A TISSUE SPECIFIC PROTEIN
; FILE REFERENCE: 129.34US01 (511582003400)
; CURRENT APPLICATION NUMBER: US/09799,250B
; CURRENT FILING DATE: 2003-07-14
; NUMBER OF SEQ ID NOS: 721
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 560
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-799-250B-560

Query Match          56.1%; Score 23; DB 5; Length 9;
Best Local Similarity 57.1%; Pred. No. 1.2e+05;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAAQERR 8
Db 3 LSAEKR 9
```

```
Db      2 LSAEKR 8

RESULT 7
US-09-627-600C-43
; Sequence 43, Application US/09627600C
; GENERAL INFORMATION:
; APPLICANT: Isaacs, John T.
; APPLICANT: Denmeade, Samuel R.
; APPLICANT: Lilja, Hans
; TITLE OF INVENTION: ACTIVATION OF PEPTIDE PRODRUGS BY HK2
; FILE REFERENCE: 07265-191001
; CURRENT APPLICATION NUMBER: US/09/627,600C
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/146,316
; PRIOR FILING DATE: 1999-07-29
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-627-600C-43

Query Match      51.2%; Score 21; DB 5; Length 7;
Best Local Similarity 66.7%; Pred. No. 1.2e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      4 AOERVV 9
        ||:||||
Db      2 AQKRL 7

RESULT 8
US-09-627-600D-43
; Sequence 43, Application US/09627600D
; GENERAL INFORMATION:
; APPLICANT: Isaacs, John T.
; APPLICANT: Denmeade, Samuel R.
; APPLICANT: Lilja, Hans
; TITLE OF INVENTION: ACTIVATION OF PEPTIDE PRODRUGS BY HK2
; FILE REFERENCE: 07265-191001
; CURRENT APPLICATION NUMBER: US/09/627,600D
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/146,316
; PRIOR FILING DATE: 1999-07-29
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-627-600D-43

Query Match      51.2%; Score 21; DB 5; Length 7;
Best Local Similarity 66.7%; Pred. No. 1.2e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      4 AOERVV 9
        ||:||||
Db      2 AQKRL 7

RESULT 9
US-09-627-600C-93
; Sequence 93, Application US/09627600C
; GENERAL INFORMATION:
; APPLICANT: Isaacs, John T.
; APPLICANT: Denmeade, Samuel R.
; APPLICANT: Lilja, Hans
; TITLE OF INVENTION: ACTIVATION OF PEPTIDE PRODRUGS BY HK2
; FILE REFERENCE: 07265-191001
; CURRENT APPLICATION NUMBER: US/09/627,600C
```

```
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/146,316
; PRIOR FILING DATE: 1999-07-29
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 93
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-627-600C-93

Query Match      51.2%; Score 21; DB 5; Length 8;
Best Local Similarity 66.7%; Pred. No. 1.2e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      4 AOERVV 9
        ||:||||
Db      3 AQKRL 8

RESULT 10
US-09-627-600D-93
; Sequence 93, Application US/09627600D
; GENERAL INFORMATION:
; APPLICANT: Isaacs, John T.
; APPLICANT: Denmeade, Samuel R.
; APPLICANT: Lilja, Hans
; TITLE OF INVENTION: ACTIVATION OF PEPTIDE PRODRUGS BY HK2
; FILE REFERENCE: 07265-191001
; CURRENT APPLICATION NUMBER: US/09/627,600D
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/146,316
; PRIOR FILING DATE: 1999-07-29
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 93
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-627-600D-93

Query Match      51.2%; Score 21; DB 5; Length 8;
Best Local Similarity 66.7%; Pred. No. 1.2e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      4 AOERVV 9
        ||:||||
Db      3 AQKRL 8

RESULT 11
US-09-627-600C-124
; Sequence 124, Application US/09627600C
; GENERAL INFORMATION:
; APPLICANT: Isaacs, John T.
; APPLICANT: Denmeade, Samuel R.
; APPLICANT: Lilja, Hans
; TITLE OF INVENTION: ACTIVATION OF PEPTIDE PRODRUGS BY HK2
; FILE REFERENCE: 07265-191001
; CURRENT APPLICATION NUMBER: US/09/627,600C
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/146,316
; PRIOR FILING DATE: 1999-07-29
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 124
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-627-600C-124

Query Match      48.8%; Score 20; DB 5; Length 5;
Best Local Similarity 80.0%; Pred. No. 1.2e+05;
```

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 AQERR 8  
||:|  
Db 1 AQRR 5

## RESULT 12

US-09-627-600D-124  
; Sequence 124, Application US/09627600D  
; GENERAL INFORMATION:  
; APPLICANT: Isaacs, John T.  
; APPLICANT: Denmeade, Samuel R.  
; APPLICANT: Lilja, Hans  
; TITLE OF INVENTION: ACTIVATION OF PEPTIDE PRODRUGS BY HK2  
; FILE REFERENCE: 07265-191001  
; CURRENT APPLICATION NUMBER: US/09/627,600D  
; CURRENT FILING DATE: 2000-07-28  
; PRIOR APPLICATION NUMBER: US 60/146,316  
; PRIOR FILING DATE: 1999-07-29  
; NUMBER OF SEQ ID NOS: 135  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 124  
; LENGTH: 5  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-627-600D-124

Query Match 48.8%; Score 20; DB 5; Length 5;  
Best Local Similarity 80.0%; Pred. No. 1.2e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 AQERR 8  
||:|  
Db 1 AQRR 5

## RESULT 13

US-09-390-061D-251  
; Sequence 251, Application US/09390061D  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Livingston, Brian D.  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Baker, Denise Marie  
; APPLICANT: Celis, Esteban  
; APPLICANT: Kubo, Ralph T.  
; APPLICANT: Grey, Howard M.  
; APPLICANT: Epimmune, Inc.  
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Plasmodium falciparum  
; FILE REFERENCE: 2060.0040004  
; CURRENT APPLICATION NUMBER: US/09/390,061D  
; CURRENT FILING DATE: 1999-09-03  
; PRIOR FILING DATE: 1998-02-03  
; PRIOR APPLICATION NUMBER: US 08/821,739  
; PRIOR FILING DATE: 1997-03-20  
; PRIOR APPLICATION NUMBER: US 08/753,615  
; PRIOR FILING DATE: 1996-11-23  
; PRIOR APPLICATION NUMBER: US 60/013,833  
; PRIOR FILING DATE: 1996-03-21  
; PRIOR APPLICATION NUMBER: US 08/590,298  
; PRIOR FILING DATE: 1996-01-23  
; PRIOR APPLICATION NUMBER: US 08/452,843  
; PRIOR FILING DATE: 1995-05-30  
; PRIOR APPLICATION NUMBER: US 08/454,033  
; PRIOR FILING DATE: 1995-05-26  
; PRIOR APPLICATION NUMBER: US 08/451,913  
; PRIOR FILING DATE: 1995-05-26  
; PRIOR APPLICATION NUMBER: US 08/344,824  
; PRIOR FILING DATE: 1994-11-23  
; PRIOR APPLICATION NUMBER: US 08/278,634  
; NUMBER OF SEQ ID NOS: 3842  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 251  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Synthetic peptide  
US-09-390-061D-251

Query Match 48.8%; Score 20; DB 5; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.2e+05;

; PRIOR FILING DATE: 1994-11-23  
; PRIOR APPLICATION NUMBER: US 08/278,634  
; PRIOR FILING DATE: 1994-07-21  
; NUMBER OF SEQ ID NOS: 3842  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 251  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Synthetic Peptide  
US-09-390-061D-251

Query Match 48.8%; Score 20; DB 5; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.2e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 QERR 8  
||||  
Db 4 QERR 7

## RESULT 14

US-09-390-061D-2053  
; Sequence 2053, Application US/09390061D  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Livingston, Brian D.  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Baker, Denise Marie  
; APPLICANT: Celis, Esteban  
; APPLICANT: Kubo, Ralph T.  
; APPLICANT: Grey, Howard M.  
; APPLICANT: Epimmune, Inc.  
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Plasmodium falciparum  
; FILE REFERENCE: 2060.0040004  
; CURRENT APPLICATION NUMBER: US/09/390,061D  
; CURRENT FILING DATE: 1999-09-03  
; PRIOR FILING DATE: 1998-02-03  
; PRIOR APPLICATION NUMBER: US 09/017,743  
; PRIOR FILING DATE: 1997-03-20  
; PRIOR APPLICATION NUMBER: US 08/821,739  
; PRIOR FILING DATE: 1996-11-23  
; PRIOR APPLICATION NUMBER: US 60/013,833  
; PRIOR FILING DATE: 1996-03-21  
; PRIOR APPLICATION NUMBER: US 08/590,298  
; PRIOR FILING DATE: 1996-01-23  
; PRIOR APPLICATION NUMBER: US 08/452,843  
; PRIOR FILING DATE: 1995-05-30  
; PRIOR APPLICATION NUMBER: US 08/454,033  
; PRIOR FILING DATE: 1995-05-26  
; PRIOR APPLICATION NUMBER: US 08/451,913  
; PRIOR FILING DATE: 1995-05-26  
; PRIOR APPLICATION NUMBER: US 08/344,824  
; PRIOR FILING DATE: 1994-11-23  
; PRIOR APPLICATION NUMBER: US 08/278,634  
; NUMBER OF SEQ ID NOS: 3842  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2053  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Synthetic peptide  
US-09-390-061D-2053

Query Match 48.8%; Score 20; DB 5; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.2e+05;



Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 QERR 8  
| | | |  
Db 4 QERR 7

RESULT 15  
US-09-390-061D-2240  
; Sequence 2240, Application US/09390061D  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Livingston, Brian D.  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Baker, Denise Marie  
; APPLICANT: Celis, Esteban  
; APPLICANT: Kubo, Ralph T.  
; APPLICANT: Grey, Howard M.  
; APPLICANT: Epimmune, Inc.  
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Plasmodium falciparum  
; TITLE OF INVENTION: Using Peptide and Nucleic Acid Compositions  
; FILE REFERENCE: 2060.0040004  
; CURRENT APPLICATION NUMBER: US/09/390,061D  
; CURRENT FILING DATE: 1999-09-03  
; PRIOR APPLICATION NUMBER: US 09/017,743  
; PRIOR FILING DATE: 1998-02-03  
; PRIOR APPLICATION NUMBER: US 08/821,739  
; PRIOR FILING DATE: 1997-03-20  
; PRIOR APPLICATION NUMBER: US 08/753,615  
; PRIOR FILING DATE: 1996-11-23  
; PRIOR APPLICATION NUMBER: US 60/013,833  
; PRIOR FILING DATE: 1996-03-21  
; PRIOR APPLICATION NUMBER: US 08/590,298  
; PRIOR FILING DATE: 1996-01-23  
; PRIOR APPLICATION NUMBER: US 08/452,843  
; PRIOR FILING DATE: 1995-05-30  
; PRIOR APPLICATION NUMBER: US 08/454,033  
; PRIOR FILING DATE: 1995-05-26  
; PRIOR APPLICATION NUMBER: US 08/451,913  
; PRIOR FILING DATE: 1995-05-26  
; PRIOR APPLICATION NUMBER: US 08/344,824  
; PRIOR FILING DATE: 1994-11-23  
; PRIOR APPLICATION NUMBER: US 08/278,634  
; PRIOR FILING DATE: 1994-07-21  
; NUMBER OF SEQ ID NOS: 3842  
; SOFTWARE: Patent version 3.2  
; SEQ ID NO 2240  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Synthetic Peptide  
US-09-390-061D-2240

Query Match 48.8%; Score 20; DB 5; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.2e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 QERR 8  
| | | |  
Db 5 QERR 8

Search completed: October 7, 2003, 15:00:39  
Job time : 16 secs

**THIS PAGE BLANK (USPTO)**